EXCESS

FleXible user-CEntric Energy poSitive houseS

Deliverable 3.2: EXCESS Data Management Framework

LC-EEB-03-2019

New developments in plus energy houses (IA),

Proposal No. 870157

Title	Document Version
D3.2 EXCESS Data Management Framework	1.0

Project Number	Project Acronym	Project Title
H2020-870157	EXCESS	FleXible user-CEntric Energy poSitive houseS

Contractual Delivery Date	Actual Delivery Date	Deliverable Type*-Security**
31 August 2021	6 October 2021	D-PU

*Type: P: Prototype; R: Report; D: Demonstrator; O: Other.

Security Class: PU: Public; PP: Restricted to other programme participants (including the Commission); RE: Restricted to a group defined by the consortium (including the Commission); CO: **Confidential, only for members of the consortium (including the Commission).

Responsible	Organisation	Contributing WP
Konstantinos Latanis	Suite5	WP3

Abstract

This deliverable describes the design and development of the EXCESS Data Management Platform and the elaboration of the EXCESS Common Information Model. The EXCESS Data Management Platform comprises various components that enable through their operation the ingestion and management of data in the EXCESS system, so that they can be made available for analysis by the other EXCESS ICT components towards the realization of the PEB concept in the buildings of the four demo sites of the EXCESS project. The methodology for the creation of the EXCESS Common Information Model is provided along with its concepts and the functionalities of the EXCESS Data Management Platform are described, while a navigation to the different components of the EXCESS Data Management Platform is presented.

Keywords

EXCESS Data Management Platform, EXCESS Common Information Model, concepts, functionalities, technologies

Revision	Date	Description	Author (Organisation)
V0.5	23.09.2021	*Ready for internal review*	Konstantinos Latanis, Tasos Tsitsanis, Spiros Kousouris (S5)
V0.6	29.09.2021	*Internal review completed*	Koen Allaerts (VITO), Veikko Ikonen (VTT)
V1.0	6.10.2021	*Final version submitted to EC*	Konstantinos Latanis, Tasos Tsitsanis, Spiros Kousouris (S5)

This project has received funding from the
European Union's Horizon 2020 Research and Innovation Programme
under Grant Agreement № 870157.
More information available at www.positive-energy-buildings.eu

Copyright Statement

EXECUTIVE SUMMARY

The deliverable D3.2 “EXCESS Data Management Framework” documents the design and development of the first release of the EXCESS Data Management Platform in the context of the Work Package 3 “Technology and User Integration via ICT”. This deliverable comprises a direct outcome of the Task 3.2 “Interoperable Data Management Framework”, which aims at the design and implementation of a message-oriented-middleware infrastructure, to facilitate the information exchange between all components of the EXCESS system along with the creation of the EXCESS Common Information Model that will enable the semantic and syntactic interoperability of data in the EXCESS system.

The EXCESS Data Management Platform is responsible for the collection, pre-processing and storage of data coming from the Distributed Information Systems of the four demo sites in order to be used subsequently for analysis and visualization purposes by the other components of the EXCESS system, namely the Data Analytics Framework, the Visualization and Blockchain applications and the Model Predictive Control component, towards the realization of the PEB concept in the buildings of the four demo sites.

The EXCESS Data Management Platform comprises several different components that enable the ingestion, pre-processing and storage of data:

a) the Data Collection component, which enables the ingestion of data through various ways, such as file uploading, acquisition through APIs and Pub/Sub messaging.
b) the Data Mapping component, which offers the mechanisms for the matching of the ingested data elements to the concepts of the EXCESS Common Information Model, assuring in that way the homogeneity of data in the EXCESS system so that they can be suitable for further processing by the various EXCESS ICT components. In particular, the EXCESS Common Information Model has been elaborated based on the study of the most important standards in the building and energy domain and on the sample datasets provided by the demo site partners.
c) the Data Cleaning component, which allows the performance of cleaning rules on the collected datasets so that any erroneous data are curated.
d) the Data Anonymization component, which provides the mechanisms for protecting the privacy and anonymity of sensitive and personal data in the EXCESS Data Management Platform.
e) the Data Storage component, which enables the storage of the collected and pre-processed data in the secure repositories of the EXCESS Data Management Platform.
f) the User Management Service, which provides the necessary user registration, authentication and authorization mechanisms, defining the access rights of the users and denying any unauthorized use of data in the EXCESS Data Management Platform.

The deliverable D3.2 describes the first release of the EXCESS Data Management Platform and has received input from the deliverable D3.1 “EXCESS ICT Architecture Blueprint” regarding the technical specifications of the designed platform. This deliverable also provides input to the tasks T3.3 “Core ICT platform services”, T3.4 “Flexibility analysis and forecasting component”, T3.5 “Building/ District monitoring and control component” and T3.6 “Block chain-enabled applications for local energy communities and flexibility trading”, and their corresponding deliverables D3.3 “EXCESS Flexibility

Analytics Module", D3.4 "EXCESS Model-Predictive Control Algorithms" and D3.5 "EXCESS Blockchain Infrastructure and Applications", which are related with the design and development of the Data Analytics Framework, the MPC component and the Visualizations and Blockchain applications that will use the data stored in the EXCESS Data Management Platform for their operations. Moreover, the deliverable D3.2 will provide input to Task 4.2 "Demonstration Case Studies in main EU climatic zones" for the operation of the EXCESS Data Management Platform in the 4 demo sites of the EXCESS project. An updated version of the deliverable D3.2 will be available in M32 of the project, describing the final release of the EXCESS Data Management Platform accommodating the feedback coming from the demo sites operation and presenting any improvements and updates.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Glossary

Acronym	Full name
API	Application Programming Interface
BIM	Building Information Modelling
CIM	Common Information Model
CSV	Comma-separated Values
DoA	Description of Action
DMP	Data Management Platform
DX.Y	Deliverable X.Y
EXCESS	FlEXible user-CEntric Energy poSitive houseS
gbXML	Green Building schema
ICT	Information, Communication and Technology
IFC	Industry Foundation Classes
JSON	JavaScript Object Notation
MPC	Model Predictive Control
obXML	Occupant behaviour XML
OpenADR	Open Automated Demand Response
PEB(s)	Positive Energy Building(s)
SAREF	Smart Applications REFerence ontology
SAREF4BLDG	Smart Applications REFerence ontology for Buildings
SAREF4ENER	Smart Applications REFerence ontology for Energy
USEF	Universal Smart Energy Framework
WP	Work Package
XML	Extensible Markup Language

1 Introduction

1.1 Purpose and scope of the document

The deliverable D3.2 “EXCESS Data Management Framework” documents the outcome of the activities performed in the context of the Task 3.2 “Interoperable Data Management Framework” leading to the design and implementation of the EXCESS Data Management Platform and the EXCESS Common Information Model.

The EXCESS Data Management Platform has been designed based on the technical specifications described in the deliverable D3.1 “EXCESS ICT Architecture Blueprint” and its first implemented release enables the collection, pre-processing and storage of data so that they can be available for analysis and visualization purposes to the rest of the EXCESS ICT components towards the realization of the PEB concept in the buildings of the four demo sites of the EXCESS project.

In order to assure that interoperable and aligned data will reside in the EXCESS Data Management Platform, the EXCESS Common Information Model has been elaborated, comprising a common language for all data that are collected and stored in the EXCESS Data Management Platform. The EXCESS Common Information Model is based on the most prominent standards on the energy and building domain and is created according to the sample datasets provided by the demo site partners. These samples are based on the datasets that will be sent by the Distributed Information Systems of the four demo sites to the EXCESS Data Management Platform.

The EXCESS Data Management Platform enables the collection, pre-processing and storage of data coming from the Distributed Information Systems of the four demo sites and comprises various components for the performance of its operations, namely the Data Collection component, the Data Mapping component, the Data Cleaning component, the Data Anonymization component and the Data Storage component. The functionalities and technical details of these components are described within this deliverable and the navigation to the different components of the EXCESS Data Management Platform is provided.

The deliverable D3.2 provides input to the rest of the design and development tasks of WP3, which are T3.3, T3.4, T3.5, T3.6, and their corresponding deliverables D3.3, D3.4 and D3.5, as their related EXCESS ICT components will use the data stored in the EXCESS Data Management Platform for their operations, as well as to T4.2 that is related with the operation of the EXCESS system in the demo sites. In M32, an updated version of the D3.2 deliverable will be delivered, describing the final release of the EXCESS Data Management Platform that will encapsulate the feedback coming from the demonstrators and will include refinements and updated functionalities.

Suite5 has developed the EXCESS Data Management Platform and the EXCESS Common Information Model, while the demo site leaders are developing the interfaces of the Distributed Information Systems of the demo sites and have provided support and knowledge for the development of the EXCESS Common Information Model.

1.2 Structure of the document

In order to address all the aspects relevant to the scope of T3.2, the present deliverable has been structured as follows:

- Section 1 introduces the work performed and the scope of this deliverable along with the deliverable's structure.
- Section 2 presents an overview of the EXCESS Data Management Platform.
- Section 3 presents the methodology followed towards the definition of the EXCESS Common Information Model, along with the analysis and description of standards, ontologies and semantic data models considered as relevant to the domains-of-interest of EXCESS. Additionally, the basic definitions and terminology of the delivered data model are presented, along with its design guidelines and any decisions taken during its modelling phase.
- Section 4 describes the Data Collection component along with the technologies exploited and the related API and software information.
- Section 5 presents the Data Mapping component along with the technologies exploited and the related API and software information.
- Section 6 describes the Data Cleaning component along with the technologies exploited and the related API and software information.
- Section 7 presents the Data Anonymization component along with the technologies exploited and the related API and software information.
- Section 8 describes the Data Storage component along with the technologies exploited and the related API and software information.
- Section 9 presents the User Management service along with the technologies exploited and the related API and software information.
- Section 10 provides a thorough navigation among the different components of the EXCESS Data Management Platform.
- Finally, in section 11, the main conclusions of the work are reported.

2 EXCESS Data Management Platform Overview

The EXCESS Data Management Platform comprises a part of the EXCESS system, which shall enable through its operation the realization of the PEB concept in the buildings of the four demo sites of the EXCESS project. The EXCESS Data Management Platform constitutes the cornerstone of the EXCESS ICT Architecture as it enables the collection and management of data coming from the sensors, devices and energy components of the Distributed Information Systems in the four demo sites. The Distributed Information System of each demo site building will provide data regarding the energy demand, energy generation, energy storage and energy components operation along with building's ambient indoor conditions and weather conditions (e.g. both indoor and outdoor temperature, humidity, luminance, etc.).

The different components of the EXCESS Data Management Platform facilitate the ingestion and pre-processing of these data so that they can be subsequently used by the EXCESS Data Analytics Framework for analysis purposes, by the Model Predictive Control component for optimization of control strategies and by the visualization and blockchain applications for their operations.

As depicted in the below figure that shows the whole EXCESS high-level ICT system architecture, the EXCESS Data Management Platform comprises the Data Collection component, the Data Mapping component, the Data Cleaning component, the Data Anonymization component, the Data Storage component and the User Management service.

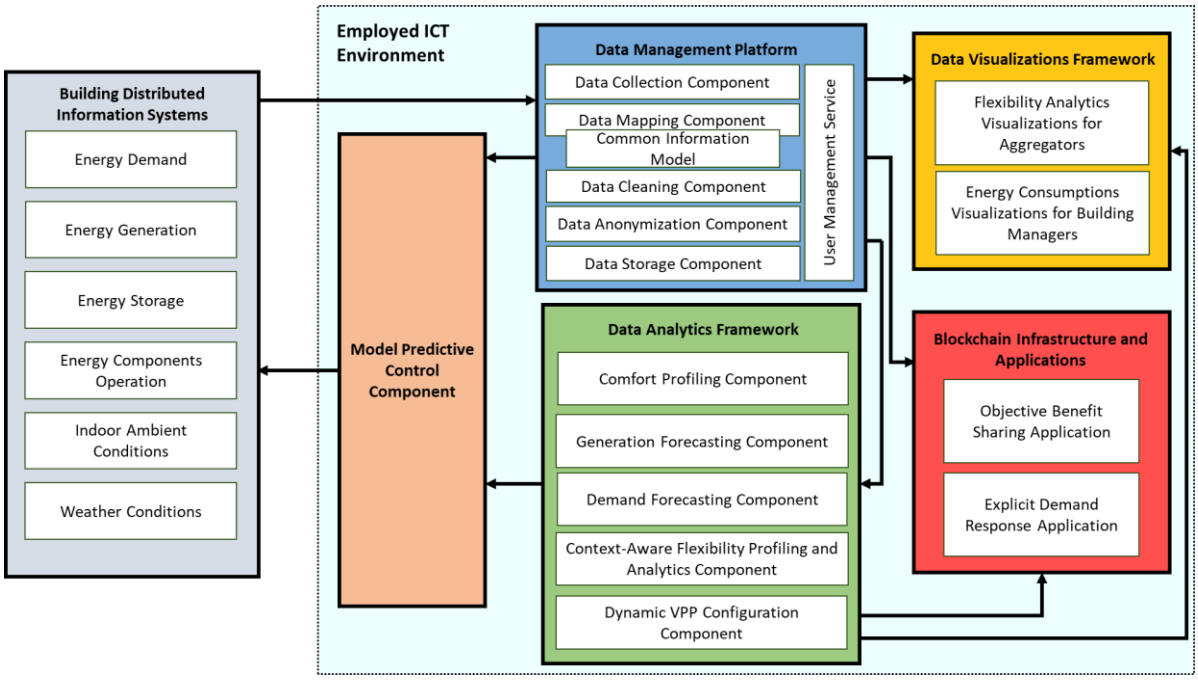

Figure 2-1: EXCESS High-Level ICT Architecture

The **Data Collection component** allows the ingestion of data coming from different sources and with various formats, such as weather data, intra-building conditions data and devices' operation data. This is enabled through the provision of the following data collection capabilities: (i) file uploading, (ii) acquisition through APIs offered by the local data platforms of the demo sites and (iii) Pub/Sub messaging.

The **Data Mapping component** enables the matching of the elements of the collected datasets to the equivalent concepts of the **EXCESS Common Information Model** in order to enhance the

interoperability and integration of the collected data in the EXCESS Data Management Platform. The **EXCESS Common Information Model** has been developed based on the most important standards of the energy and building domain and taking into account the sample datasets offered by the demo sites partners. In this sense, it comprises a common dictionary for all datasets in the EXCESS Data Management Platform that facilitates their use for further analysis and visualization purposes.

The **Data Cleaning component** enables the curation of the ingested datasets in case they have any erroneous data or inconsistencies. Therefore, the data can be cleaned, for example, from any outliers or missing values and become suitable for analysis purposes.

The **Data Anonymization component** allows the performance of mechanisms that ensure the privacy and security of personal and sensitive data coming in the EXCESS system.

The **Data Storage component** realizes the secure storing of collected datasets, after the aforementioned pre-processing activities, in the EXCESS Data Management Platform, so that they can be available for the operations of the rest of the EXCESS ICT components, namely the EXCESS Data Analytics Framework, the MPC component, the EXCESS Data Visualizations Framework and the EXCESS Blockchain Applications.

The **User Management Service** organizes the authentication and authorization mechanisms of the EXCESS Data Management Platform, specifying the access rights of the users and prohibiting any unauthorized access on data.

The first release of the EXCESS Data Management Platform is deployed at: **https://excess.s5labs.eu/** (credentials can be provided upon request)

The various components of the EXCESS Data Management Platform along with the EXCESS Common Information Model are described in further detail in the following sections of the deliverable.

3 EXCESS Common Information Model

In order to fulfil the objectives of the project, all datasets collected from the various Distributed Information Systems of the demo sites will be stored in the EXCESS Data Management Platform, under a "common language" enabling both syntactic and semantic interoperability in order to make them available for further analysis and visualization purposes in the EXCESS system.

Such a common language will be accomplished through the design and development of the EXCESS Common Information Model (CIM) constituting the cornerstone of the Data Mapping component (see Section 5), since the elements of the collected datasets will be mapped against the equivalent concepts of the EXCESS CIM resulting in a homogeneous form of data stored in the EXCESS Data Management Platform.

3.1 Methodology

Towards the definition and delivery of a thorough data model capable of addressing all the project's needs, the methodology shown below was followed, with the ultimate target to ensure that the delivered model:

- addresses all the project's demo sites data needs and end users' requirements,
- is designed based on the requirements of the EXCESS Data Management Platform,
- is built upon consideration of the most prominent standards in the building and energy domain.
- is scalable during its lifecycle, i.e., it is designed in such a way to allow for future modification, addition or deletion of its entities if required.

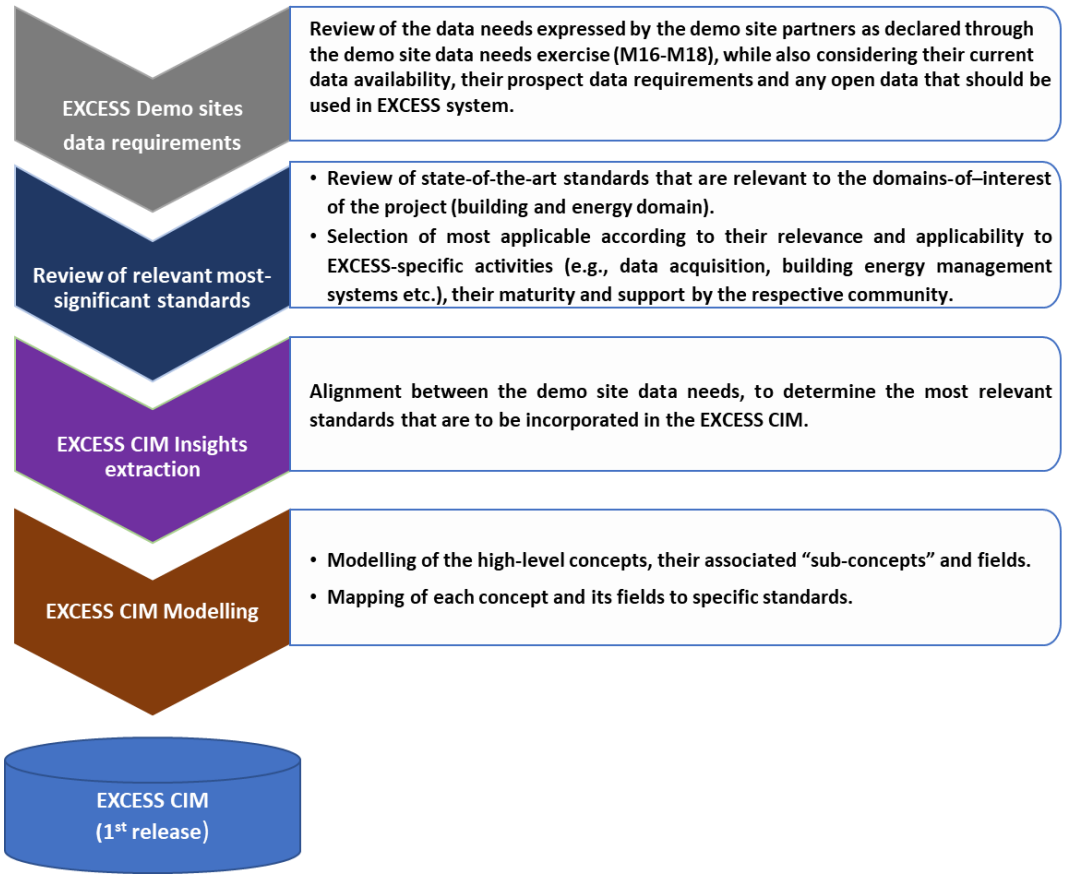

Figure 3-1: EXCESS CIM design methodology

As shown in the above figure, the initial step towards the design and definition of the first release of the EXCESS CIM, focused on the identification of the project's data requirements and needs. Through the use of an adequate template, the demo site partners declared their data availability, their prospect data needs and any open data requirements that the EXCESS system shall incorporate. This exercise made it possible to define an extensive list of all available (and future) datasets used within the EXCESS project, forming the basis of the modelling activities. The input of the demo site partners that filled in the data availability template is presented in Annex: Data Collection Exercise.

Having identified the data needs of the project's demo site partners, the focus was then given on reviewing/analysis of state-of-the-art standards considered as relevant to the EXCESS project's domains-of-interest (such as the building domain and the energy domain). Prior to the actual analysis, the selection of the most applicable material took place according to their relevance and applicability to EXCESS-specific activities (e.g., data acquisition, building energy management systems etc.), their maturity and support by the respective community. An overview of the standards reviewed is presented in the section 3.2.

The next step of the work involved the extraction of the high-level concepts that are related to the scope of EXCESS, along with their relationships (i.e., the nesting of the data model concepts) as presented in section 3.5. As an outcome of the aforementioned activities, the detailed modelling of the high-level concepts took place, along with their associated "sub-concepts" and fields. During this last step, the mapping of each concept and its fields to specific standards (where applicable) and configuration of any additional metadata that were required took place.

Overall, the EXCESS CIM has been designed and developed upon a detailed study and analysis of the data that the Distribution Information Systems of the EXCESS demo sites will attain through the installed sensors, submeters and actuators along with all deployed devices and energy systems residing in the demo site buildings. The EXCESS CIM is also based on widely-accepted standards in the energy and building domain, as described in the following section.

3.2 Related energy and building domain standards

Following the methodology for the design and development of the EXCESS CIM, this section provides an overview of the standards that are considered as highly relevant to the building and energy domain and which have been used as a basis, where applicable, during the development of the first release of the EXCESS CIM.

- **Smart Applications REFerence (SAREF) ontology**

The SAREF ontology was developed by the Dutch institute TNO aiming to facilitate interoperability between different solutions developed by different providers and among different assets in the smart applications domain [1]. The SAREF ontology aims to provide discrete and reusable elements of the ontology based on the users' needs and is built on the following key principles:

- the concepts of an existing asset can be reused and aligned accordingly
- the different elements of the ontology can be separated and/or recombined as per user needs
- the ontology can be further extended
- the processes of updating, identifying and correcting defects in the ontology, are easily maintainable [2]

The SAREF ontology enables users to create various device and technology abstraction layers and their corresponding common Application Programming Interfaces (APIs), without having knowledge of specific standards. An overview of the core concepts (i.e., main classes) of the latest version SAREF-v3.1.1 and their relationships is presented in the figure below.

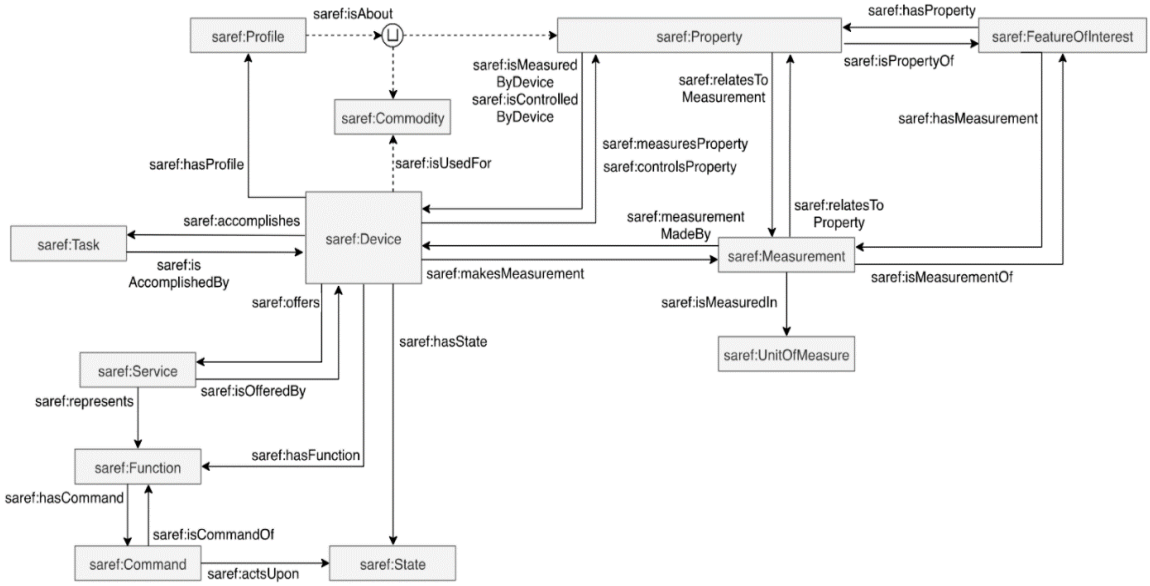

Figure 3-2: Core classes of SAREF ontology [1]

The core classes of SAREF include concepts, such as: Device (e.g., Light Switch, temperature sensor, etc.), Property (e.g., Temperature, Energy, Occupancy, etc.), Command (e.g., OnCommand, OffCommand, etc.). In general, SAREF is built in a modular way, allowing the definition of any device from pre-defined building blocks, based on the function(s) that the device performs. This can be easily seen in the following example and the above figure: A saref:Device has at least one function (saref:hasFunction). Moreover, a saref:Device can be used for (saref:isUsedFor property) offering a

commodity, such as saref:Water or saref:Gas. It can also measure a property, such as saref:Temperature, saref:Energy and saref:Smoke. Moreover, a device may consist of other devices (saref:consistsOf property) [1].

Further to the core SAREF ontology, more extensions have been published for different domains (e.g., Energy, Building, Environment, Smart Cities, Smart Agriculture, Water domain, etc.). The extensions considered to be highly relevant to the EXCESS CIM design include SAREF4ENER that focuses on the energy domain and SAREF4BLDG for the building domain, both analysed in this section.

- **SAREF4ENER**

SAREF4ENER constitutes an extension of the SAREF ontology, expressed as an OWL-DL ontology. It was developed through the collaboration of a key Italy (Energy@Home[1]) and Germany (EEBus[2]) based industry association towards enabling the interconnection of their different data models. The latest version SAREF4ENER v1.1.2 extends the core SAREF ontology with 63 additional classes, 17 object properties and 40 data type properties [1]. In general, SAREF4ENER supports interoperability among different products/services developed in the smart home domain (e.g., smart appliances) from manufacturers that support the Energy@Home or EEBus data models with the focus given on demand response scenarios, where customers can offer flexibility to the Smart Grid through management of their smart home devices via any energy management system both at home or remote [3].

- **Open Automated Demand Response (OpenADR)**

Published by the OpenADR Alliance in 2010, the OpenADR is an internationally adopted smart-grid data model that enables the information exchange related to Demand Response (DR) programs between Electricity Service Providers, Aggregators and Consumers. OpenADR enables the management of the various distributed energy resources (DER) for flexibility providers, such as Aggregators and Utility companies. In general, the OpenADR is used to structure the messages exchanged between the different stakeholders involved in automatic demand response (Auto-DR) and DER management activities, in a consistent and interoperable way. OpenADR cannot be considered as a communication protocol per se, since it relies on existing open standards such as XML for exchanging DR messages and reports. The OpenADR was developed to automate and simplify DR and DER management activities through dynamic price and reliability signals allowing electricity consumers to regulate their energy usage, save money and improve their energy efficiency, resulting in improving the overall effectiveness of power distribution across the smart grid [4]. Recently, the International Electrotechnical Commission (IEC) has approved the OpenADR 2.0 Profile Specification as a Publicly Available Specification (PAS), meaning that the OpenADR will become an IEC international standard.

- **Universal Smart Energy Framework (USEF)**

The USEF is an international standard facilitating the integration of various smart energy services and products and promoting energy flexibility trading by defining the various roles and their interactions, so that the demand-side participation can be fully utilized [5]. Developed by the USEF Foundation[3], a non-profit industry association, this standard defines a market-oriented framework for standardizing the energy flexibility trading without constraining how the trading should be implemented, accepting both bilateral and exchange-based trading.

[1] http://www.energy-home.it
[2] http://www.eebus.org/en
[3] https://www.usef.energy/usef-foundation/

Presently, in its latest version, USEF v1.3.6 is considered as the most important standard controlling the various energy flexibility trading market mechanisms. Each of the roles defined in USEF and their duties can be mapped to their real-life application in a local market; a brief description of the main roles defined in USEF is provided as follows:

- A Balance Responsible Party (BRP) is responsible for delivering supply and demand balance and identifying strategies that can reduce cost for covering potential network imbalances.
- A Distribution System Operator (DSO) is responsible for the optimal operation of the overall distribution network; depending on the market, a DSO might also carry out BRP responsibilities.
- An Aggregator, responsible for the management of the accumulated energy flexibility provided by the prosumers and based on the requirements set by the BRP.
- A Common Reference Operator (CRO) is responsible for assigning the congestion points and congestions to other involved parties.
- A Meter Data Company responsible for collecting and validating the energy consumers' metering data.
- An Active Demand and Supply is defined as the various entities that can be actively controlled with appropriate signals to adjust the energy supply and demand.
- A Prosumer who is essentially the end user that can both consume and produce energy.

- **Industry Foundation Classes (IFC)**

The IFC, published by buildingSMART International, offers a digital representation of the building domain and its associated assets, enabling information exchange among the various AEC stakeholders involved in a particular project or building asset over its lifecycle. The latest version, IFC4 is provided as an open specification for Building Information Modelling (BIM) data and comprises terms, concepts and data specification items that are derived from use within disciplines, trades, and professions that are involved in the building industry [6]. IFC4 includes the data schema, and reference data that are represented as EXPRESS or as an XML schema specification and as XML definitions of property and quantity definitions, respectively.

The IFC specification comprises four different conceptual layers namely: a) the Resource layer, including all individual schemas containing the definitions of the resources, b) the Core layer, which includes the kernel schema and the core extension schemas that contain the most general entity definitions, c) the Interoperability layer containing entity definitions specific to general products, processes or resource specialization that are used across various disciplines and d) the Domain layer, representing the higher-level layer containing entity definitions that are specializations of products, processes or resources that are related to a certain discipline, and they can be exchanged or shared within the whole data schema [6].

- **SAREF4BLDG**

SAREF4BLDG[4] constitutes an extension of the SAREF ontology based on the IFC standard for building information exchanges. It includes devices defined by the IFC4 and intends to represent devices and other physical objects typically found in buildings, thus enabling an efficient interaction and integration among actors (e.g., engineers, architects, consultants, contractors, component manufacturers, etc.), processes and tools, during the life cycle of a building (Planning and Design,

[4] https://saref.etsi.org/saref4bldg/v1.1.2/

Construction, Commissioning, Operation, Retrofitting/Refurbishment/Reconfiguration, and Demolition/Recycling). SAREF4BLDG is delivered as an OWL-DL ontology that extends SAREF with 72 classes, 179 object properties and 83 data type properties [7]. Such an ontology provides mechanisms to facilitate the exchange and interoperability of data between actors involved in various stages of a building life cycle [3]. SAREF4BLDG has been identified as closely relevant to EXCESS, since it has been created upon consideration of the overall Architecture, Engineering and Construction (AEC) domain. The core classes of its latest version, SAREF4BLDG v1.1.2, are shown in the figure below.

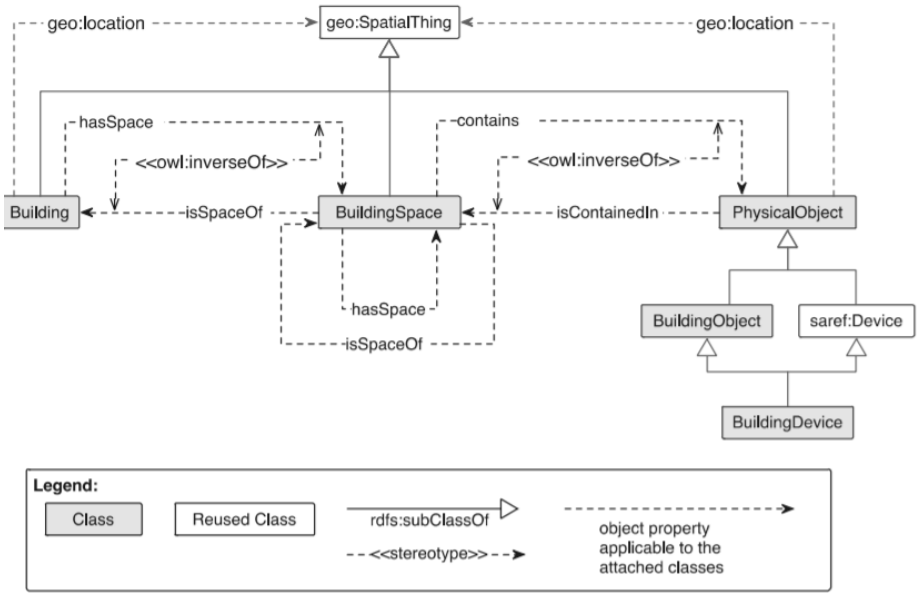

Figure 3-3: Overview of the SAREF4BLDG extension [7]

- **Green Building schema (gbXML)**

The gbXML has been created with the purpose to enable interoperability between 3D building information models (BIM) and architectural/engineering analysis software such as AutoCAD, ArchiCAD, etc., enabling them to communicate by carrying out the transfer of building information through an XML language schema. gbXML has been the most commonly used industry standard schema for over 20 years, and it is currently supported by 58 building energy analysis, BIM authoring and CAD software tools of different vendors [8]. The latest gbXML schema (v.6.01) is available from the website[5], and includes over 500 elements and attributes that can be related to a building description. gbXML can be followed to define the relevant data exchange/storage formats for building information, increasing the overall data exchange efficiency. Usage of gbXML enables storage of detailed building data (geometry, constructions, heating, cooling, air flows, lightning, etc.).

- **Occupant behaviour XML (obXML)**

On the way to address the need for a common language and new modelling processes for representing building occupant's behaviour, the IEA Annex 66 developed the DNAs framework, titled obXML schema and expressed as an XML schema. The main scope of obXML is to facilitate the linking between three main elements representing buildings, occupants and their behaviour [9]. The main part of the work is the behaviours' part which models occupant behaviour on the basis of Drivers, Needs, Actions

[5] https://www.gbxml.org/

and Systems. In EXCESS CIM, such information can be utilised to define the relevant data exchange/storage formats for building information, increasing the overall data exchange efficiency.

3.3 Main outcomes and challenges

Having provided an overview of the standards identified as highly relevant to the scope of EXCESS CIM, this section presents a summary of the main findings and challenges that shall be taken into consideration for the design and development of the EXCESS CIM.

- There is no existing standard addressing in full the requirements of the intended EXCESS CIM and that could be adopted directly as the project's internal model.
- Many of the reviewed material define strict ontological relations; which means that users have to adhere to the selection of abstractions set in the standards. Any additions in the ontology will require an adjustment of the overall ontological definition and placement of a new class within the modelling ontology. This is unfavourable for the case of EXCESS CIM, as it is foreseen to expand and evolve during the project's lifecycle.
- One of the key challenges to be addressed is the different level of development in different fields; i.e., many users utilise the IFC extensively for building data model representation, while others use ontologies defining a building asset as an operating system (e.g., SAREF4BLDG). It is thus evident that accomplishing interoperability with these standards is essential.
- Even though, there are existing standards that cover quite extensively domain/fields relevant to the EXCESS (such as the energy and building domain), they have a certain focus; which represents a challenge since the EXCESS data modelling shall be applicable in different fields.

3.4 EXCESS CIM design considerations and guidelines

Prior to presenting the design consideration and guidelines driving the overall development of the EXCESS CIM, this section provides initially the definition of its basic terms:

- **Concept**: An abstraction of a real-world/physical object (e.g., Sensing Device, Building) in the EXCESS CIM that describes its characteristics (e.g., id, type, etc.), and the relationship with other concepts, under a particular domain. For example, a Sensing Device could be modelled as a concept, with its properties/attributes (such as sensor measurement) defined as its fields.
- **Field**: A single characteristic/ property of a certain concept modelled in the EXCESS CIM (e.g., for the "SensingDevice" concept, potential fields might include its id, manufacturer, name, type, measurementUnit, serialNumber, measurementDateTime, status, etc.). The relations between the various concepts are modelled by defining an object-type field that refers to another defined concept. In the case of a "SensingDevice" example, the relationship between the SensingDevice and its measurements is defined via an object-type field "relatedSensingDeviceMeasurements" which refers to the concept measurement that includes fields such as id, description, value, etc.

A number of design considerations that have been made during the development of the EXCESS CIM are presented, constituting also functional requirements for the data model. Such requirements entail what is needed from the delivered data model, in order to fulfil its functionalities and successfully support the operations of the EXCESS Data Management Platform (DMP).

1. As the EXCESS CIM will be responsible for effectively managing the semantics of the data collected in the EXCESS Data Management Platform while ensuring interoperability, the EXCESS CIM shall be able to incorporate the proper level of semantics enabling interoperability, without making it difficult to implement it such as by following very-strict standards.
2. The EXCESS CIM shall support the efficient operation of the EXCESS Data Management Platform by ensuring that all the related semantics are kept together with the collected data in order to facilitate the analysis and visualization purposes of the EXCESS system.
3. In order for EXCESS to fulfil its objectives successfully, interoperability must be accomplished both at technical and at semantic level. It is thus essential to deliver consistent and non-ambiguous data interpretation towards properly identifying the semantics; which represents a key challenge for a system that interacts with several different data sources.
4. In order for EXCESS to be able to integrate existing data from diverse data sources, the EXCESS CIM will require the addition of new modelling properties which aim directly at enabling interoperability.
5. The EXCESS CIM shall be designed to allow extensibility by providing adequate extension processes and allowing model upgrades, as the EXCESS system may potentially evolve.
6. The EXCESS CIM shall utilise existing standards in order to provide mappings to existing data representations resulting from the respective data collection tasks.

Further to the analysis of relevant domain-specific data standards and the identification of the EXCESS-specific data needs, the delivery of the EXCESS CIM relies also on the definition of an appropriate design, adhering to particular guidelines. Such guidelines (see table below) were decided based on the aforementioned considerations and drove the overall design activities of the EXCESS CIM, resulting in the definition of its high-level domain concepts and their respective fields and related concepts.

Table 1: EXCESS CIM design guidelines

ID	Description
01	The EXCESS CIM shall include all concepts and their relationships that could be modelled in terms of the data, in order to cover both the current data needs of the EXCESS end users, but also future data needs that may arise.
02	All the concepts and their relationships in the EXCESS CIM shall consist of an adequate number of fields that describe effectively (as practicably possible) a particular real-world object.
03	The EXCESS CIM shall comply with existing relevant data standards focusing on the energy and building data exchange modelling. Nevertheless, adoption of their concepts and fields does not imply that the design of the EXCESS CIM shall focus on a specific direction but rather towards addressing the EXCESS end users' needs.
04	The EXCESS CIM shall be updated and assessed on a regular basis to preserve consistency and ensure its efficient evolution. As such, the EXCESS CIM shall deliver functionalities including addition, update, and deprecation of its concepts/fields ensuring extension and scalability.
05	The level of abstraction between concepts, fields, and relations shall be properly defined, since the data to be loaded in the EXCESS Data Management Platform by the demo site partners might have a different structure.
06	The granularity level of the data that are to be loaded in the EXCESS Data Management Platform shall be also considered to determine whether the data denote static or dynamic properties of a concept or field.
07	The EXCESS CIM shall adhere to the following naming and syntactic rules:

	○ The names of the concepts and fields may consist of word combinations as long as their meaning is easily understood. ○ Concepts' names shall comply with the upper camel case convention, i.e. without spaces or punctuation, indicating the separation of words with a single capitalized letter and the first letter of the entire word is uppercase. ○ Fields' names shall comply with the lower camel case convention, i.e. the first letter of the entire word is lowercase, but subsequent first letters are uppercase. ○ The names of concepts and fields shall be expressed as nouns.
08	The description of concepts and fields shall be clear and easy to understand, while a common domain vocabulary shall be adopted to avoid the use of alternative words that describe the same concepts, in order to ensure a common interpretation.
09	The EXCESS CIM shall include clear and explicit concepts; meaning that the level of concept nesting shall be chosen based on the specific case, so as to maintain only clear connections between the related concepts.
10	The various data types to be supported by the EXCESS CIM are listed as follows: • Date (e.g., 2021-06-01) • Datetime (e.g., 2021-06-01 13:33:00) • Time (e.g., 18:22:45) • Double (e.g., 0.003) • Integer (e.g., 12) • Boolean (e.g., TRUE or FALSE) • Object (implies a reference to another concept) • String (e.g., "text")
11	Depending on the data type, the values of data shall follow a specific format or be within a specific range, allowing the appropriate transformations to take place in the EXCESS Data Management Platform. For example, the values of "double" data type shall include also their measurement unit (e.g., km, m, cm etc.), while the values of "datetime" data type should follow a particular form and their corresponding time zone (e.g., UTC).

3.5 Definition of EXCESS CIM high-level concepts

Taking into consideration the analysis of the EXCESS project's requirements and that of the landscaping of standards in the previous sections, as well as the aforementioned design considerations and guidelines, the definition of the EXCESS CIM high-level concepts is presented in the table below, specifying their category (mainly for their presentation needs) and defining the associated fields of each high-level concept.

Overall, this first version of the EXCESS CIM includes 92 high-level concepts along with their fields. It shall be noted that the resulting EXCESS CIM is designed and modelled to be as complete as possible, including all possible concepts and fields. Nevertheless, this first version of the EXCESS CIM is expected to be updated with additions (or deprecations) of concepts/fields so as to address all data needs emerging throughout the project's implementation, where necessary.

Table 2 EXCESS CIM: High level concepts and associated fields

High-level Concept	***Fields***
Category: Generic	
Address	id, name, description, type, addressLine, cityCode, cityName, countryCode, countryName, districtName, streetName, buildingName, buildingNumber, floorNumber, apartmentID, plotID, postalCode, postOfficeBox, region, status
ContactPerson	name, type, description, title, familyName, role, emailAddress, departmentName, jobTitle, contactNumber, faxNumber
Location	id, name, description, type, altitude, latitude, longitude, latitudeDirectionIndicator, systemID, radius longitudeDirectionIndicator,PolygonPoint
Period	id, name, description, type, startDateTime, endDateTime, earliestStartDateTime, latestEndDateTime, maxDuration, minDuration, defaultDuration, durationAbsoluteUncertainty, durationPercentUncertainty, optionalIndicator, pauseTime, referenceDateTime, referenceDay, referenceDayNumber, referenceMonth referenceMonthNumber, referenceYear, referenceWeekNumber, seasonCode, sequenceNumeric, weekdayIndicator, weekendIndicator
Status	id, name, description, type, manualOverride, conditionCode, conditionIndicator,referenceDateTime, reason, reasonCode, remarks, statusCount, testResult, overchargedStatusIndicator, underchargedStatusIndicator
Category: Equipment	
AirConditioner	id, name, description, type, airflowRate, brandName, code, model, coolingCapacity, coolingMedium, externalSurfaceArea, europeanSeasonalEnergyEfficiencyRatio, heatingCapacity, internalRefrigerantVolume, internalSurfaceArea, internalWaterVolume, manufacturerName, maximumCapacity, minimumCapacity, nominalCapacity, nominalCoolingLoad, nominalHeatingLoad, nominalVoltage, nominalHeatTransferArea, nominalHeatTransferCoefficient, nominalNoiseLevel, nominalPartLoadMax, nominalPartLoadMin, nominalPartLoadRatio, performanceCoefficient, refrigerantName, refrigerantClass, serialNumber seasonalEnergyEfficiencyRatio, supplyAirTemperature, exitAirTemperature, wasteAirTemperature
AirConditionerControlOperation	id, name, description, type, coolingSetpoint, createdDateTime, updatedDateTime, fanDirection, fanSpeed, heatingSetpoint, timerIndicator, temperatureSetpoint,modeSetting, powerSwitch, reportedDateTime
AutomatedOperationProfile	id, name, description, activatedDateTime, configuration, createdDateTime, updatedDateTime, status, remoteControlIndicator, reselectionSupportIndicator, singleSlotSchedulingIndicator, totalSequencesCount

Battery	id, name, description, type, code, duration, autonomy, brandName, model, serialNumber, cellsInParallelCount, cellsInSeriesCount, chargeVoltage, cutOffVoltage, cycleLife, depthOfDischarge, dischargeRate, manufacturerName, maxCapacity, maxChargeCurrent, maxChargeRate, maxDischargeCurrent, minChargeStatus, nominalCapacity, nominalEnergyDelivered, nominalVoltage, ratedCapacity, roundTripEfficiency
BatteryControlOperation	id, name, description, type, chargingSwitch, createdDateTime, updatedDateTime, reportedDateTime, dischargingSwitch, stateOfChargeIndicator
Boiler	id, name, description, type, brandName, code, serialNumber, energySource, model, flowMode, heatingLoad, heatingSurfaceArea,manufacturerName, operatingMode, nominalCapacity, nominalEnergyConsumption, nominalEnergyConsumptionRate, nominalPartLoadRatio, nominalPower, nominalThermalEfficiencyoutletTemperatureMax, outletTemperatureMin, outletTemperatureRange, parasiticElectricConsumption, parasiticElectricLoad, pressureRating, storageCapacityAvailability, waterInletTemperatureMax, waterInletTemperatureMin, waterInletTemperatureRange, waterStorageCapacity
BoilerControlOperation	id, name, description, type, powerSwitch, createdDateTime, reportedDateTime, temperatureSetpoint, timerIndicator, updatedDateTime
Device	description, type, serialNumber, brandName, code, deviceName, expectedPowerType, hardwareRevision, manufacturerLabel, manufacturerName, model, nodeID, nominalPower, powerSkewness, powerSource, powerStandardDeviation, powerMax, powerMin, softwareRevision, vendorCode, vendorName
DeviceControlEvent	id, type, createdDateTime, occurenceDateTime, occurenceDescription, reportedDateTime, updatedDateTime
DeviceControlEventOperation	id, name, description, type, capacityPercent, setpoint, consumeActionDescription, createdDateTime, updatedDateTime, reportedDateTime,offsetLevel, offsetPercent, operationModeChange, produceActionDescription
DeviceControlStatus	id, name, description, type, createdDateTime, reportedDateTime, updatedDateTime, consumeLoadControlStatus produceLoadControlStatus,
ElectricVehicle	id, name, description, type, brandName, code, efficiency, manufacturerName, nominalDrivingRange, nominalVoltage, releaseDateTime, serialNumber, typicalRechargeTime, upstreamEmissions
ElectricVehicleChargingPoint	id, name, type, chargingMode, effectiveChargingPower, effectiveCurrent, maxChargingPower, maxCurrent
ElectricVehicleChargingPointControlOperation	id, name, description, type, chargingModeSwitch, chargingPowerClass, chargingTime, createdDateTime, updatedDateTime, powerSwitch, reportedDateTime, timeIndicator
DomesticWaterHeater	id, name, description, type, brandName, code, manufacturerName, model, title, nominalPerformanceEffciency, serialNumber, nominalPower, volume
DomesticWaterHeaterControlOperation	id, name, description, type, timerIndicator powerSwitch,createdDateTime, updatedDateTime, reportedDateTime, temperatureSetpoint
Gateway	id, name, description, type, brandName, code, manufacturerName, model, serialNumber

BufferTank	id, name, description, type, serialNumber, brandName, code, storageCapacity, storageLevel, storageMaxTemperature
BufferTankControlOperation	id, name, description, type, timerIndicator powerSwitch,createdDateTime, updatedDateTime, reportedDateTime, temperatureSetpoint
BoreholeThermalEnergyStorageSystem	id, name, description, type,flowRate, supplyTemperature, returnTemperature, HeatOutput
BoreholeThermalEnergyStorageSystemControlOperation	id, name, description, type, timerIndicator powerSwitch,createdDateTime, updatedDateTime, reportedDateTime
HeatPump	id, name, description, type, serialNumber, brandName, code, airflowRate, energyEfficiencyRatio, heatingCapacity, manufacturerName, maximumCapacity, minimumCapacity, model, nominalCapacity, nominalHeatingLoad, nominalHeatTransferCoefficient, nominalVoltage, performanceCoefficient, outgoingTemperature, returnTemperature, mode, electricalConsumption, heatingOutput, coolingOutput, condensorSupplyTemperature, condensorReturnTemperature, condensorFlowrate, energyConsumption, evaporatorSupplyTemperature, evaporatorReturnTemperature, evaporatorFlowrate, thermalOutput
HeatPumpControlOperation	id, name, description, type, createdDateTime, heatingSetpoint, powerSwitch, reportedDateTime, updatedDateTime, temperatureSetpoint, timerIndicator
LightingDevice	id, name, description, brandName, code, serialNumber, model, colorAppearance, colorRenderingIndex, colorTemperature, contributedLuminousFlux, lampBallastType, lampCompensationType, lampMaintenanceFactor, manufacturerName, nominalPower, numberOfDimmingScales, spectrumMin, spectrumMax, spectrumRange
LightingDeviceControlOperation	id, name, description, type, colorSetting, colorTemperatureSetting, createdDateTime, dimmingLevel, modeSetting, powerSwitch, reportedDateTime, timerIndicator, updatedDateTime
MeteringSystem	id, name, description, type, acquiredDateTime, brandName, code, manufacturerName, model, serialNumber
SensingDevice	id, name, description, type, serialNumber accuracy, brandName, code, manufacturerName, model
SmartAppliance	id, name, description, type, brandName, code, manufacturerName, model, nominalVoltage, serialNumber, stateOfCharge, typicalEnergyConsumption
SmartApplianceControlOperation	id, name, description, type, heatingSetpoint, powerSwitch, createdDateTime, reportedDateTime, updatedDateTime, temperatureSetpoint, timerIndicator
SpaceHeater	id, name, description, type, bodyMass, brandName, code, energySource, heatTransferDimension, heatTransferMedium, manufacturerName, model, mode, nominalPower, nominalVoltage, outputCapacity, panelsCount, placementType, sectionsCount, serialNumber, temperatureClassification, thermalEfficiency, thermalMass, typicalEnergyConsumption

SpaceHeaterControlOperation	id, name, description, type, heatingSetpoint, powerSwitch, createdDateTime, reportedDateTime, updatedDateTime, temperatureSetpoint, timerIndicator
VentilationSystem	id, name, description, type, brandName, code, serialNumber, defrostIndicator, heatTransferType, manufacturerName, model, operationTemperatureMax, operationTemperatureMin, primaryAirFlowRateMax, primaryAirFlowRateMin, secondaryAirFlowRateMax, secondaryAirFlowRateMin
VentilationSystemControlOperation	id, name, description, type, createdDateTime, fanSpeed, powerSwitch, reportedDateTime, updatedDateTime, timerIndicator
Category: Measurements	
EnergyConsumptionMeasurements	id, name, description, type, airConditionerLoad, baseLoad, batteryLoad, boilerLoad, createdDateTime, updatedDateTime, deviceLoad, diversifiedLoad, diversityFactor, peakLoad, forecastDateTime, forecastLoad, gridLoad, load, lightingDeviceLoad, loadFactor, loadProfileHourly, observedDateTime, smartApplianceLoad, EVChargingPoint spaceHeaterDeviceLoad, totalConsumptionHourly, totalConsumption, unmetLoad, utilizationFactor, ventilationSystemLoad
EnergyProductionMeasurements	id, name, description, type, activePower, frequency, activeEnergyExport, totalUptime, createdDateTime, updatedDateTime, observedDateTime , activeEnergyExportGeothermal, activeEnergyExportPhotovoltaic, activeEnergyExportRenewable, activeEnergyExportSolar, activeEnergyExportSolarThermal, apparentPower, activeEnergyExportWind, averageCurrent, equivalentAvailabilityFactor, forecastDateTime, grossGeneration, grossGenerationGeothermal, grossGenerationPhotovoltaic, grossGenerationRenewable, grossGenerationSolar, grossGenerationSolarThermal, grossGenerationWind, netCapacityFactor, netCapacityFactorGeothermal, netCapacityFactorPhotovoltaic, netCapacityFactorRenewable, netCapacityFactorSolar, netCapacityFactorSolarThermal, netCapacityFactorWind, netGeneration, netGenerationGeothermal, netGenerationPhotovoltaic, netGenerationRenewable, netGenerationSolar, netGenerationSolarThermal, netGenerationWind, peakCurrent, peakToPeakCurrent, peakToPeakVoltage, peakVoltage, powerFactor, rmsCurrent, rmsVoltage, totalEnergyExport, totalEnergyExportGeothermal, totalEnergyExportPhotovoltaic, totalEnergyExportRenewable, totalEnergyExportSolar, totalEnergyExportSolarThermal, totalEnergyExportWind, totalEnergyImport, totalPrimaryEnergySupply, totalPrimaryEnergySupplyPhotovoltaic, totalPrimaryEnergySupplyGeothermal, totalPrimaryEnergySupplyRenewable, totalPrimaryEnergySupplySolar, totalPrimaryEnergySupplySolarThermal, totalPrimaryEnergySupplyWind
EnergyStorageMeasurements	id, name, description, actualEnergyStored, chargeSetpoint, chargeStatus, createdDateTime, energyDeliveredSinceLastCharge, energyObtainedFromStorage, forecastDateTime, numberofCycles, observedDateTime, operationTimeSinceLastCharge, remainingUsefulLife, stateOfCharge, stateOfHealth, stateOfSafety, targetEnergyStored, throughput, totalEnergyDelivered, totalOperationTime, updatedDateTime
Measurement	id, name, description, value, createdDateTime, measuredDateTime, measurementUnit, reportedDateTime

SensingDeviceMeasurements	acousticPressure, airQualityIndex, alarmTemperature, batteryCoolantIntakeTemperature, createdDateTime, observedDateTime, batteryCollantOutputTemperature, forecastMaxTemperature, vocConcentration co2Concentration, forecastMinTemperature, forecastTemperature, forecastTemperatureRange, maxTemperature, meanTemperature, minTemperature, noiseLevel, occurenceDateTime, observedLuminance, observedSoundPower, observedSoundPowerLevel, observedTemperature, temperatureChangeRate, waterFlowRate
MeteorologicalMeasurements	id, absoluteHumidity, atmosphericPressure, cloudiness, conditionDetails, conditionIntensity, conditionStatus, description, measuredDateTime, precipitationProbability, precipitationRate, proximity, realFeelTemperature, relativeHumidity, reportedDateTime, seaLevelPressureMax, seaLevelPressureMean, seaLevelPressureMin, temperature, temperatureMax, temperatureMin, turbulenceIntensity, uvIndex, visibilityMax, visibilityMean, visibilityMin, windDirection, windspeed, solarIrradiation, totalSolarRadiance, directSolarRadiance solarRadiationHorizontal, solarRadiationDiffuse, solarRadiationVerticalWest, solarRadiationVerticalSouth
Category: Plant	
PhotovoltaicGenerator	id, name, description, type, brandName, capitalCost, code, model, serialNumber, deratingFactor, efficiency, levelisedCostOfElectricity, lifetime, manufacturerName, maxPowerPoint, maxPowerPointCurrent, maxPowerPointVoltage, nominalCapacity, nominalVoltage, openCircuitVoltage, operatingCost, panelsCount, powerTolerance, replacementCost, shortCircuitCurrent, temperatureCoefficient, temperatureCoefficientPercent, supplyPipetemperature, returnPipetemperature
PowerPlant	id, name, description, type, activeIndicator, code, energySource, site, status, typicalPowerOutput
RenewableGenerator	id, name, description, controlMode, deadband, discreteMode, phase, status, targetUnit, targetValue
VirtualPowerPlant	id, name, description, type, capacity, energySource, powerSourcesCount, scalability
MeteorologicalStation	id, name, description, type, manufacturerName, model
WindGenerator	id, type, additionalCapacity, cutBackInWindSpeed, cutInWindSpeed, cutOutWindSpeed, lifetime, noiseLevel, nominalCapacity, nominalWindSpeed, operationTemperatureMax, operationTemperatureMin, operationTemperatureRange, survivalSpeed, type, typicalEfficiency
GeothermalPlant	id, name, description, type, activeIndicator, code, energySource, site, status, typicalPowerOutput
Category: KPI	
KeyPerformanceIndicator	id, name, description, type, code, measurementUnit
KeyPerformanceIndicatorValue	measurementUnit, createdDateTime, maxValue, minValue, referenceDateTime, relativeDeviation, tolerance, updatedDateTime, value, valueRange, absoluteDeviation

Category: Network	
ACLine	id, name, description, type, capacity, cableType, conductorDiameter, coreDiameter, impedance, layersCount, length, name, nominalCurrent, nominalCurrentMax, nominalFrequency, nominalVoltage, operatingCurrentMax, operatingVoltageMax, reactance, resistance, shortCircuitTemperature, shuntConductance, shuntSusceptance, status, type, transformersCount, zeroReactance, zeroResistance, zeroShuntConductance, zeroShuntSusceptance
ConnectivityNode	id, name, connectionID, groundID, nodeType, nominalActivePower, nominalReactivePower, nominalVoltage, phaseType
Grid	id, name, type, baseLoad, description, firmCapacity, nominalCapacity, nominalFrequency, operatingReserve, operatingReservePercent, peakDemand, spinningReserve, spinningReservePercent
Category: Flexibility	
AggregatorPortfolio	id, name, description, appliancesCount, availabilityTimeline, connection, createdDate, generationAssetsCount, powerCapacity, prosumersCount, storageDevicesCount, updatedDate
DemandSideManagementEvent	id, name, description, type, code, comment, createdDateTime, description, id, marketContext, status, priority, modificationDateTime, modificationCount, modificationReason, name, occurenceDateTime, startDateTime, rampUpDuration, testEvent
DemandSideManagementEventSignal	id, name, type, currentValue, level, measurementUnit, targetValue, customerBidEnergySetpoint, energyPrice, customerBidLoadSetpoint, customerBidPrice, demandChargePrice, demandChargePriceRelative, demandChargePriceMultiplier, electricityPrice, electricityPriceRelative, electricityPriceMultiplier, energyPriceMultiplier, energyPriceRelative
DemandSideManagementReportReading	id, name, description, type, accuracy, confidence, marketContext, periodicSamplingStatus, periodicSamplingIndicator
DemandSideManagementReport	id, name, description, type, backDuration, code, createdDateTime, dataSource, status, duration, granularity, marketContext, optInType, requestID
Flexibility	name, type, activationTime, baselineFlexibility, demandFlexbility, defaultDuration, duration, flexibilityForecast, forecastDateTime, generationFlexibility, observedDateTime, price, storageFlexibility
LoadResponse	id, name, status, exponentModel, pConstantCurrent, pConstantImpedance, pConstantPower, pFrequencyExponent, qConstantCurrent, qConstantImpedance, qConstantPower, qFrequencyExponent
Category: Flexibility Market	
FlexibilityContract	id, name, description, type, duration, activationDateTime, capacity, startDate createdDate, updatedDate, effectiveDate, expiryDate, customerType, payment, flexibiliyEventsNumber, penalties, phase, price, serviceCategory, terms, signatureDate

FlexibilityContractPricingStructure	id, name, description, type, code, status, region, customerType, baselineFlexibility, ceilingUsageConsumption, startDate, createdDate, endDate, ceilingUsagePower flexibilityPrice, floorUsageConsumption, floorUsagePower, serviceCategory
FlexibilityMarket	id, name, description, type, duration, ancillaryService, balanceForecast, balancingRequirements, reason, balancingService, clearanceDateTime, closureDateTime, constraint, demandForecast, price, primaryReserve, flexibilityRequirements, generationForecast, marketRole, marketVolume, participant, referenceDateTime, secondaryReserve, tertiaryReserve, settlementDateTime
FlexibilityOffer	id, name, description, baselineReference, creationDateTime, expiryDateTime,
FlexibilityOfferOption	Price, duration, activationDateTime, activationTimes, maxFlexibilityCapacity, minActivationFactor, optionReference
FlexibilityRequest	id, name, description, type, activationDateTime, creationDateTime, duration, flexibilityAmount, status
FlexibilitySettlement	id, actualPower, availablePower, baselineReference, deliveredFlexibilityPower, netSettlement, offeredPower, orderedFlexibilityPower, penalty, powerDeficiency, price, requestedPower, reservedPower, settlementDate, settlementStatus
EnergyTariffProfile	id, name, description, chargeKind, code, consumptionSequenceNumber, consumptionStartValue, creationDate, endDateTime, fixedCostPart, sequenceNumber, scheduleCycle, startDateTime, variableCostPart
Category: Building	
Building	id, name, description, type, subType, actualGrossArea, actualNetArea, bimFile, constructionEndDate, constructionMethod, constructionStartDate, coolingWetBulb, documentation, eavesHeight, totalHeight, elevationOfRefHeight, elevationOfTerrain, energyPerformanceCertificationClass, fireProtectionClass, grossFloorArea, heatingDesignDateTime, heatingDryBulb, heatingWetBulb, landmarkIndicator, netFloorArea,, lastRefurbishmentDate, occupancyType, permanentIndicator, plannedGrossArea, plannedNetArea, planningControlStatus, zonesCount smartReadinessAssessmentClass, sprinklerProtectionAutomaticIndicator, sprinklerProtectionStatus, storeysCount, heatDemandForecast, aggregatedEnergybalance
BuildingZone	id, name, type, actualGrossArea, actualNetArea, documentation, plannedGrossArea, plannedNetArea, spacesCount, floorNumber
BuildingSpace	id, name, description, documentation, elevation, elevationWithFlooring, energyPerformanceCertificationClass, finishCeilingHeight, finishFloorHeight, grossFloorArea, netFloorArea, occupancyType, spaceHeight, smartReadinessAssessmentClass
BuildingFloor	id, name, description, actualGrossArea, actualNetArea, documentation, elevation, grossHeight, netHeight, plannedGrossArea, plannedNetArea, floorNumber, zonesCount
HumanComfort	id, name, description, type, preference, acousticValueMax, acousticValueMin, createdDateTime, feedback, IAQValueMax, IAQValueMin, occurenceDateTime, optimalAcousticValue, optimalIAQValue,

	optimalThermalValue, optimalVisualValue, optimalAcousticRange, optimalIAQRange, optimalThermalRange, optimalVisualRange, satisfactionIndicator, thermalValueMin, thermalValueMax, visualValueMin, visualValueMax
BuildingOccupancy	id, name, description, type, areaPerOccupant, code, maxOccupantsCount, minOccupantsCount, occupancyTimePerDay, occupancyTimePerWeekDay, occupancyTimePerWeekendDay, occupantsCount, occupantsPeakCount
BuildingOccupant	id, name, description, age, title, birthDateTime, category, familyName, gender, givenName, lifestyle, maidenName, middleName, nationality, profession,
Category: Stakeholders	
DemandSideAggregator	id, name, description, type, brandName, departmentName, legalClassificationCode, legalName
BuildingManager	id, name, description, type, brandName, departmentName, legalClassificationCode, legalName
LocalAdministration	id, name, description, type, legalName, departmentName, legalClassificationCode
DistributionNetworkOperator	id, name, description, type, brandName, departmentName, legalName
FlexibilityMarketOperator	id, name, description, type, brandName, departmentName, legalClassificationCode, legalName
EnergyServiceProvider	id, name, description, type, brandName, departmentName, legalClassificationCode, legalName
LocalEnergyCommunity	id, name, description, type, brandName, departmentName, legalClassificationCode, legalName, members, membersCount
Prosumer	id, name, description, age, birthDateTime, category, familyName, gender, givenName, maidenName, middleName, nationality, profession, title, lifestyle
PowerPlantOperator	id, name, description, type, brandName, departmentName, legalClassificationCode, legalName
RenewableGeneratorOperator	id, name, description, type, brandName, departmentName, legalClassificationCode, legalName
Category: Incidents	
Event	id, name, description, type, createdDateTime, modificationDateTime, modificationReason, occurenceDateTime, priority, remarks, testIndicator, updatedDateTime
Incident	id, name, description, type, subType actualRestorationTime, cause, effect, priority, status, estimatedRestorationTime, forecastDateTime, workNotes, occurenceDateTime
IncidentLog	id, name, description, type, subType,createdDateTime, updatedDateTime, totalFailureCount
Schedule	id, name, description, type, activeIndicator, createdDateTime, duration, plannedDateTime, plannedMaintenanceTime
TroubleTicket	id, name, description, type, code, createdDateTime, reportedDateTime, resolvedDateTime, responsibleWorker
Event	id, name, description, type, createdDateTime, modificationDateTime, modificationReason, occurenceDateTime, priority, remarks, testIndicator, updatedDateTime

3.6 Alignment of EXCESS CIM concepts to relevant standards

As previously mentioned, the concepts of the EXCESS CIM were extracted upon analysis of existing relevant standards/models. Under this context and towards ensuring consistency, Table 3 below presents for each of the concepts included in the EXCESS CIM, the source concept and the relevant standard(s) which they align with (where applicable).

Table 3 EXCESS CIM: High level concepts and relevant standards

EXCESS CIM Category	EXCESS CIM Concept	Source Concept	Relevant Standard
General	Address	-	N.A
	ContactPerson	-	N.A
	Period	-	N.A
	Location	-	N.A
	Status	-	N.A
Equipment	AirConditioner	AirConditioningSystem	SAREF4BLDG
	AirConditionerControlOperation	-	N.A
	AutomatedOperationProfile	AutomatedOperationProfile	SAREF4ENER
	Battery	Battery	SAREF4ENER
	BatteryControlOperation	-	N.A
	Boiler	Boiler	SAREF4BLDG
	BoilerControlOperation	-	N.A
	Device	Device	SAREF, SAREF4ENER
	DeviceControlEvent	DeviceControlEvent	SAREF4ENER
	DeviceControlEventOperation	DeviceControlEventAction	SAREF4ENER
	DeviceControlStatus	DeviceControlStatus	SAREF4ENER
	ElectricVehicle	-	N.A
	ElectricVehicleChargingPoint	-	N.A
	ElectricVehicleChargingPointControlOperation	-	N.A
	DomesticWaterHeater	-	N.A
	DomesticWaterHeaterControlOperation	-	N.A
	Gateway	-	N.A
	BufferTank	-	N.A
	BufferTankControlOperation	-	N.A
	BoreholeThermalEnergyStorageSystem	-	N.A
	BoreholeThermalEnergyStorageSystemControlOperation	-	N.A
	HeatPump	-	N.A
	HeatPumpControlOperation	-	N.A
	LightingDevice	LightingDevice	SAREF4BLDG
	LightingDeviceControlOperation	-	N.A
	MeteringSystem	MeteringSystem	SAREF
	SensingDevice	Sensor	SAREF

	SmartAppliance	SmartAppliance	SAREF4BLDG
	SmartApplianceControlOperation	-	N.A
	SpaceHeater	SpaceHeatingDevice	SAREF4BLDG
	SpaceHeaterControlOperation	-	N.A
	VentilationSystem	VentilationSystem	SAREF4BLDG
	VentilationSystemControlOperation	-	N.A
Measurements	EnergyConsumptionMeasurements	-	N.A
	EnergyProductionMeasurements	-	N.A
	EnergyStorageMeasurements	-	N.A
	Measurement	Measurement	SAREF
	SensingDeviceMeasurements	-	N.A
	MeteorologicalMeasurements	-	N.A
Plant	PhotovoltaicGenerator	-	N.A
	PowerPlant	-	N.A
	RenewableGenerator	-	N.A
	VirtualPowerPlant	-	N.A
	MeteorologicalStation	-	N.A
	WindGenerator	-	N.A
	GeothermalPlant	-	N.A
KPI	KeyPerformanceIndicator	-	N.A
	KeyPerformanceIndicatorValue	-	N.A
Network	ACLine	-	N.A
	ConnectivityNode	-	N.A
	Grid	-	N.A
Flexibility	AggregatorPortfolio	AggregatorPortfolio	USEF
	DemandSideManagementEvent	DemandResponseEvent	OpenADR
	DemandSideManagementEventSignal	DemandResponseEvent Sign al	OpenADR
	DemandSideManagementReportReading	DemandResponseReportRea ding	OpenADR
	DemandSideManagementReport	DemandResponseReport	OpenADR
	Flexibility	Flexibility	USEF
	LoadResponse	-	N.A
Flexibility Market	FlexibilityContract	Contract	USEF
	FlexibilityContractPricingStructure	-	N.A
	FlexibilityMarket	-	N.A
	FlexibilityOffer	Offer	USEF
	FlexibilityOfferOption	OfferOption	USEF
	FlexibilityRequest	Request	USEF
	FlexibilitySettlement	Settlement	USEF
	EnergyTariffProfile	-	N.A
Building	Building	Building	IFC
	BuildingZone	BuildingZone	IFC
	BuildingSpace	BuildingSpace	IFC
	BuildingFloor	BuildingStorey	IFC
	HumanComfort	Needs	obXML
	BuildingOccupancy	Occupancy	IFC
	BuildingOccupant	-	N.A

Stakeholders	DemandSideAggregator	-	N.A
	LocalAdministration	-	N.A
	DistributionNetworkOperator	-	N.A
	FlexibilityMarketOperator	-	N.A
	EnergyServiceProvider	-	N.A
	LocalEnergyCommunity	-	N.A
	Prosumer	Prosumer	USEF
	PowerPlantOperator	-	N.A
	RenewableGeneratorOperator	-	N.A
	BuildingManager	-	N.A
Incidents	Incident	-	N.A
	IncidentLog	-	N.A
	Schedule	-	N.A
	TroubleTicket	-	N.A
	Event	-	N.A

4 Data Collection component

4.1 Design and functionalities

The Data Collection component is responsible for the ingestion of data coming from various sources and with different formats in the EXCESS Data Management Platform. The Data Collection component offers a user-friendly interface that enables the data provider to define the necessary data collection configurations. The collection of data can be realized either through uploading of files (e.g., uploading of historical data in the EXCESS Data Management Platform) or through APIs or Pub/Sub mechanisms that are offered by the distributed information systems of the demo sites.

The functionalities of the Data Collection component are described below:

a) **Definition of the data collection process in a user-friendly way:** The Data Collection component allows the configuration of the data ingestion process in the desired way through an easy-to-use interface, by enabling the selection of the collection options, the authentication specifications and other related details.

b) **Data ingestion through file uploading:** The uploading of files of different formats, such as (i) tabular (e.g. CSV, TSV), (ii) non-tabular (e.g. XML, JSON) and (iii) others (i.e. non-text data) is enabled. The uploading of data samples of files is also provided, which drives the creation of the configuration files, including the data collection details defined by the data provider.

c) **Data acquisition through APIs:** The Data Collection component enables the collection of data through the APIs that are offered by the local data platforms of the demo sites by providing a comprehensive interface for the definition of the API details. Moreover, the already configured API connection is tested and in case of success, sample data are collected. In addition, the management of data update through the APIs is facilitated by configuring the corresponding scheduling details.

d) **Management of API authentication details:** The Data Collection component allows the data provider to define the authentication details of an API connection by setting the type of authentication and the related information, such as tokens or credentials. The Data Collection component uses these authentication details to verify the API connection.

e) **Data acquisition through Pub/Sub mechanisms:** The Data Collection component enables the collection of streaming data through Pub/Sub mechanisms that are offered by the local data platforms of the demo sites by providing an easy-to-use interface for the definition of the connection and retrieval settings. Also, sample data are collected upon testing of the already configured Pub/Sub mechanisms.

f) **Option for storage of specific data:** During the data collection process, the data provider is enabled to specify the part of the data sample that s/he would like to be stored in the EXCESS Data Management Platform. In that context, only the chosen data will be further processed, while the rest of them will be rejected and not be stored eventually in the EXCESS Data Management Platform.

g) **Management of data update:** Regarding file uploading, the Data Collection component enables the update of an already stored dataset by allowing the uploading of an additional file with the same data structure. In this way, the already stored dataset will be appended with

the records of the new file, using the already defined pre-processing rules of the existing configuration file to speed-up the update process.

4.2 Technologies and tools

The Data Collection component implementation is written in Python[6]. The frontend comprises the user interface of the Data Collection component and is developed in VueJS[7] and TailwindCSS[8]. The backend includes the different data collection options and is implemented with the Flask micro web framework[9]. The orchestration engine, based on Kubernetes[10], manages the containerized services of the different processes in the EXCESS Data Management Platform. The Data Collection component uses a relational database for storage purposes exploiting PostgreSQL[11], along with a data lake for temporary storage needs utilizing MinIO[12]. RESTful interfaces are utilized for the communication between the frontend and the backend of the Data Collection component using Swagger[13].

4.3 APIs information

The Data Collection component communicates with the other components of the EXCESS Data Management Platform through a messaging functionality. The communication between the frontend and the backend of the Data Collection component is realized through internal APIs which support intra-component integration.

4.4 Software package repository

The Data Collection component is closed source and no source code is available publicly. The source code and the related deployment instructions are maintained in the related private repositories and the corresponding subcomponents are containerized with Docker[14].

[6] Python, https://www.python.org/
[7] Vue.js, https://vuejs.org/
[8] TailwindCSS, https://tailwindcss.com/
[9] Flask, https://flask.palletsprojects.com/en/2.0.x/
[10] Kubernetes, https://kubernetes.io/
[11] PostgreSQL, https://www.postgresql.org/
[12] MinIO, https://min.io/
[13] Swagger, https://swagger.io/
[14] Docker, https://www.docker.com/

5 Data Mapping Component

5.1 Design and functionalities

The Data Mapping component is responsible for the mapping of the elements of the collected datasets to the equivalent concepts of the EXCESS Common Information Model, enabling in that way the elaboration of interoperable and consistent datasets that can be utilized for further analytical processes in the EXCESS system. The EXCESS Common Information Model has been constructed based on the most important data standards in the energy and building domain and by analyzing and extracting the different entities from the datasets coming from the various Distributed Information Systems of the four demo sites and will be updated as long as new elements are needed to be mapped.

The functionalities of the Data Mapping component are described below:

a) **Exploitation of various matching techniques for automated mapping predictions:** The Data Mapping component maps the data elements of the ingested dataset to the related concepts of the EXCESS Common Information Model. The automated mapping predictions are executed using different fuzzy matching techniques.

b) **Manual configuration of proposed mapping predictions:** Through a user-friendly interface the Data Mapping component enables the data provider to check the proposed automated mappings and choose whether they should be maintained, updated or deleted. In addition, any unidentified concepts can be mapped manually to related concepts of the EXCESS Common Information Model. Moreover, the data types, measurement units and any other data transformations may be specified.

c) **Intuitive exploration of the EXCESS Common Information Model:** The Data Mapping component enables the data provider to explore the Common Information Model, view its structure and get deeper knowledge of its concepts, allowing him/her in that way to choose if any manual mappings are more suitable for the data elements of his/her dataset.

5.2 Technologies and Tools

The Data Mapping component implementation is written in Python exploiting the NumPy[15], Pandas[16] and scikit-learn[17] libraries. The frontend comprises the user interface of the Data Mapping component utilizing VueJS and TailwindCSS. The backend, which is developed based on the Flask micro web framework, includes the different data mapping services. The transformation service realizes the unit and any other data transformations, the prediction manager facilitates the automatic mapping process and the mapping configuration manager enables the manual modification of mapping predictions and the storage of final mapping selections in the configuration file. The Data Mapping component uses a relational database for storage purposes based on PostgreSQL, along with a data lake for temporary storage needs exploiting MinIO and an indexing engine for the EXCESS Common Information Model using Elasticsearch[18]. RESTful interfaces, based on Swagger, are utilized for the communication between the frontend and the backend of the Data Mapping component.

[15] Numpy, https://numpy.org/
[16] Pandas, https://pandas.pydata.org/
[17] https://scikit-learn.org/stable/
[18] Elasticsearch, https://www.elastic.co/

5.3 APIs information

The Data Mapping component communicates with the other components of the EXCESS Data Management Platform through a messaging functionality. The communication between the frontend and the backend of the Data Mapping component is realized through internal APIs which support intra-component integration.

5.4 Software package repository

The Data Mapping component is closed source and no source code is available publicly. The source code and the related deployment instructions are maintained in the related private repositories and the corresponding subcomponents are containerized with Docker.

6 Data Cleaning Component

6.1 Design and functionalities

The Data Cleaning component is responsible for the performance of the necessary quality checks on the collected data and the employment of the designated by the data provider cleaning rules in order to curate these data in case they include any errors or inconsistencies. The Data Cleaning component facilitates the curation of collected and mapped datasets so that they become suitable for further analysis purposes.

The functionalities of the Data Cleaning component are described below:

a) **Definition and performance of cleaning rules:** The cleaning of the data is performed by the Data Cleaning component in case any errors are found during the validation checks, by offering a series of different cleaning rules that can be selected by the data provider in order to have curated datasets in the EXCESS Data Management Platform.
b) **Option for elimination of missing values or outliers:** In case there are missing values or outliers in the ingested dataset, the Data Cleaning component enables the data provider to drop from the dataset the rows that contain such missing values.
c) **Option for replacement of missing values or outliers:** In case there are missing values or outliers in the ingested dataset, the Data Cleaning component enables the data provider to replace these dataset fields by selecting through a list of potential replacement values, such as a specific default value set by the data provider, the maximum or the minimum value of the column where the missing value or outlier exists, etc.

6.2 Technologies and Tools

The Data Cleaning component implementation is written in Python exploiting the NumPy and Pandas libraries. The frontend comprises the user interface of the Data Cleaning component utilizing VueJS and TailwindCSS. The backend, which is developed based on the Flask micro web framework, includes the quality checking and cleaning services. In addition, the cleaning configuration manager enables the storage of cleaning selections in the configuration file. The Data Cleaning component uses a relational database for storage purposes based on PostgreSQL, along with a data lake for temporary storage needs exploiting MinIO. RESTful interfaces, based on Swagger, are utilized for the communication between the frontend and the backend of the Data Cleaning component.

6.3 APIs information

The Data Cleaning component communicates with the other components of the EXCESS Data Management Platform through a messaging functionality. The communication between the frontend and the backend of the Data Cleaning component is realized through internal APIs which support intra-component integration.

6.4 Software package repository

The Data Cleaning component is closed source and no source code is available publicly. The source code and the related deployment instructions are maintained in the related private repositories and the corresponding subcomponents are containerized with Docker.

7 Data Anonymization Component

7.1 Design and functionalities

The Data Anonymization component is responsible for the performance of anonymization rules designated by the data provider in order to ensure the privacy and anonymity of datasets in case they include any sensitive or personal data or even quasi-identifiers that may reveal a person's identity through certain combinations of them. Through a user-friendly interface, the Data Anonymization component enables the data provider to choose how s/he may handle such sensitive information in order to store such datasets in the EXCESS Data Management Platform in a protected and secure manner.

The functionalities of the Data Anonymization component are described below:

a) **Definition and performance of anonymization rules:** The Data Anonymization component provides an easy-to-use interface where the data provider can choose from a variety of different anonymization rules, which are employed in order to achieve the privacy and anonymity of datasets in the EXCESS Data Management Platform.
b) **Option for elimination of identifying data:** In case the data provider designates that a column of a dataset contains data that identify personal or private information, the Data Anonymization component performs the dropping of this column in order to ensure the privacy of the dataset in the EXCESS Data Management Platform.
c) **Option for masking or grouping of quasi-identifiers:** In case the data provider designates that a column of a dataset contains quasi-identifier data, the Data Anonymization component performs the masking or numerical grouping of this column (depending on the data type of the column, string or number respectively) at the necessary level in order to ensure the privacy of the dataset in the EXCESS Data Management Platform.

7.2 Technologies and Tools

The Data Anonymization component implementation is written in Python exploiting the NumPy and Pandas libraries. The frontend comprises the user interface of the Data Anonymization component utilizing VueJS and TailwindCSS. The backend, which is developed based on the Flask micro web framework, includes the anonymization service. In addition, the anonymization configuration manager enables the storage of anonymization selections in the configuration file. The Data Anonymization component uses a relational database for storage purposes based on PostgreSQL, along with a data lake for temporary storage needs exploiting MinIO. RESTful interfaces, based on Swagger, are utilized for the communication between the frontend and the backend of the Data Anonymization component.

7.3 APIs information

The Data Anonymization component communicates with the other components of the EXCESS Data Management Platform through a messaging functionality. The communication between the frontend and the backend of the Data Anonymization component is realized through internal APIs which support intra-component integration.

7.4 Software package repository

The Data Anonymization component is closed source and no source code is available publicly. The source code and the related deployment instructions are maintained in the related private repositories and the corresponding subcomponents are containerized with Docker.

8 Data Storage Component

8.1 Design and functionalities

The Data Storage component is responsible for storing data coming from the Distributed Information Systems of the four demo sites. As soon as the data have been collected and pre-processed, they are stored in a secure storage space in the EXCESS Data Management Platform. The Data Storage component facilitates scalability and big data management optimization, while data model indexing is used for model search performance improvement.

The functionalities of the Data Storage component are described below:

a) **Storage of data collection tasks and related configurations:** The Data Storage component stores the properties of the data collection task along with the corresponding settings of the configuration file during the data collection configuration process.
b) **Data persistence:** The Data Storage component stores the dataset along with the data sample, as soon as the data collection and pre-processing activities have been completed.
c) **Data model indexing:** Following the storage of the EXCESS Common Information Model, the Data Storage component creates the necessary indexes for the stored data model in order to facilitate faster data model searching during the data mapping process.
d) **Intermediate data storage:** The Data Storage component offers a temporary storage space, where the intermediate configuration and data files that are produced during the various data processes are stored, allowing in this way the pause and continuation of these processes. This approach enhances the fast resuming of these processes and supports traceability in case a specific data process fails.

8.2 Technologies and Tools

The Data Storage component includes for the realization of its functionalities a relational database based on PostgreSQL, a non-relational database using MongoDB[19], a temporary storage data lake based on MinIO and a data model indexing engine using Elasticsearch.

8.3 APIs information

No external APIs are used by the Data Storage component.

8.4 Software package repository

The Data Storage component is closed source and no source code is available publicly. The source code and the related deployment instructions are maintained in the related private repositories and the corresponding subcomponents are containerized with Docker.

[19] MongoDB, https://www.mongodb.com/

9 User Management Service

9.1 Design and functionalities

The User Management Service provides a series of features referring to the use of the EXCESS Data Management Platform by the users. More specifically, it offers the necessary mechanisms for user registration, authentication, login and analysis of the actions they perform over the different components involved in the EXCESS Data Management Platform, as well as it facilitates the specification of different user roles and groups in the EXCESS Data Management Platform. Moreover, the appropriate authorization and access control functionalities are applied for the verification of access rights of users over specific datasets according to the restrictions imposed by the data providers. In this sense, specific rights are assigned to the different users of the EXCESS Data Management Platform for accessing the data that are stored and are made available through the platform, while assuring that non-authorized users or user types will not have access to any data. Furthermore, the User Management Service offers a series of usage analytics that present information in a user-friendly way about the use of datasets by specific users and user groups, such as which datasets are searched or accessed more, how many datasets a user is accessing, etc., along with information and activity logs of the users over the different components and datasets (e.g. new dataset creation, modification of an existing dataset).

9.2 Technologies and Tools

The User Management service implementation is written in Python exploiting the NumPy and Pandas libraries. The frontend comprises the user interface of the User Management service utilizing VueJS. The backend, which is developed based on the Flask micro web framework, includes the user management and access policies services. RESTful interfaces, based on Swagger, are utilized for the communication between the frontend and the backend of the User Management service.

9.3 APIs information

The User Management service communicates with the other components of the EXCESS Data Management Platform through a messaging functionality. The communication between the frontend and the backend of the User Management service is realized through internal APIs which support intra-component integration.

9.4 Software package repository

The User Management service is closed source and no source code is available publicly. The source code and the related deployment instructions are maintained in the related private repositories and the corresponding subcomponents are containerized with Docker.

10 Navigation to the EXCESS Data Management Platform

Within this section, the navigation to the EXCESS Data Management Platform across its different components will be presented and the various functionalities of the EXCESS Data Management Platform will be displayed through descriptive screenshots.

10.1 Data Collection

In the first page of the EXCESS Data Management Platform, the user (i.e. data provider) sees the list of the already created data collection jobs along with their details, such as the execution status of each job and the steps that are included in each job. In addition, the user has the option to edit or delete a data collection job.

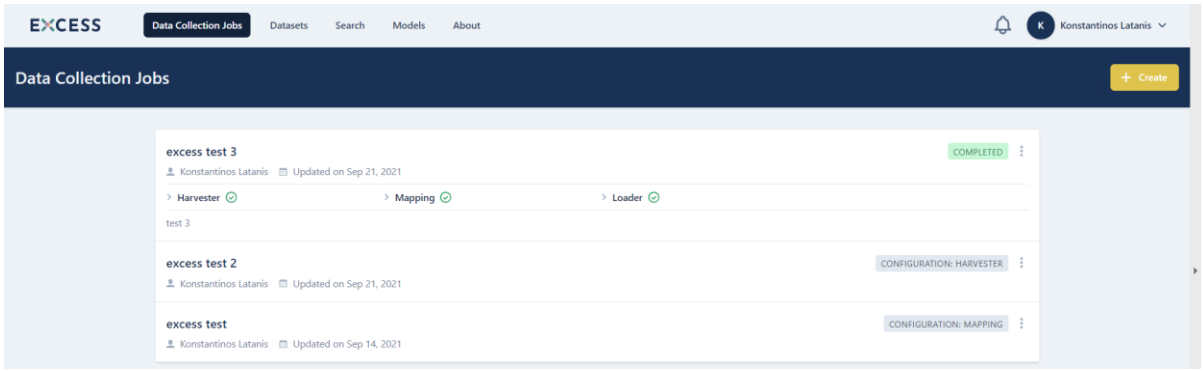

Figure 10-1: List of data collection jobs

By clicking on the "Create" button, the user can define the name and the description of a new data collection job along with the pre-processing steps that it will include. The harvesting and loading steps are mandatory. If the mapping step is selected for processing text data, the cleaning and anonymizing steps are also made available. For uploading files as single objects, such as images, the mapping step -and eventually the cleaning and anonymization steps- shall not be selected.

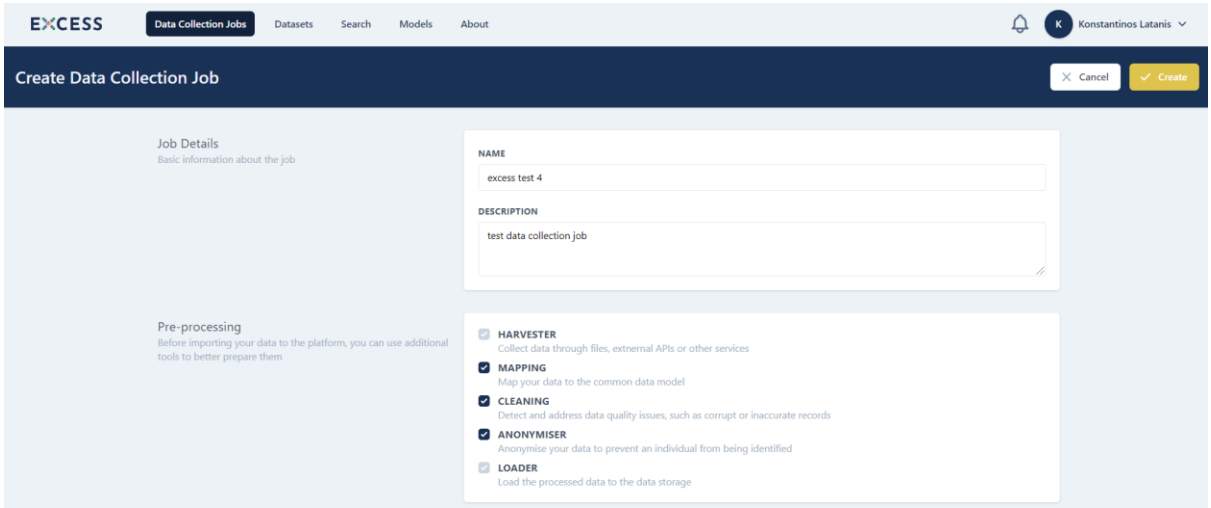

Figure 10-2: Create a data collection job

After creating the new data collection job, which is displayed in the aforementioned list of data collection jobs, the user may click on the data collection step where s/he is prompted to choose the method of data collection that can be:

- File uploading
- Acquisition through APIs offered by the local data platforms of the demo sites
- Pub/Sub mechanisms available by the local data platforms of the demo sites for streaming data

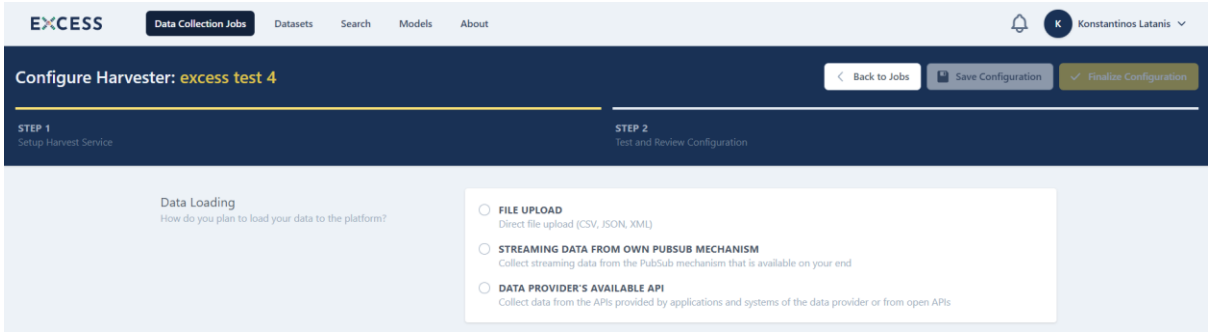

Figure 10-3: Selection of data collection method

10.1.1 File uploading

When the user chooses to upload a file (e.g. uploading of historical data in EXCESS Data Management Platform to facilitate analysis purposes), s/he has to define the format of the file (i.e. CSV, XML, JSON) and also upload a data sample and the full file. The data sample contains a few rows of the file and is used in order to set-up quickly the pre-processing rules of the data collection job that are stored in the configuration file and can also facilitate the future update of the file to be uploaded.

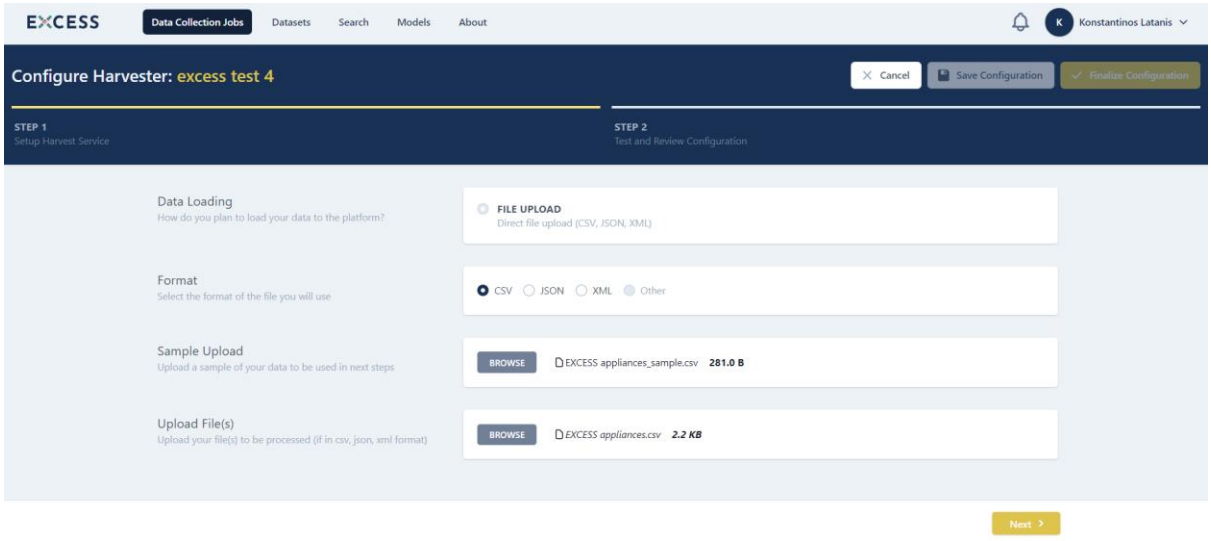

Figure 10-4: Configure file uploading options

By clicking on the “Next” button, the user is enabled to check the data sample before the next configuration activities of the EXCESS Data Management Platform begin.

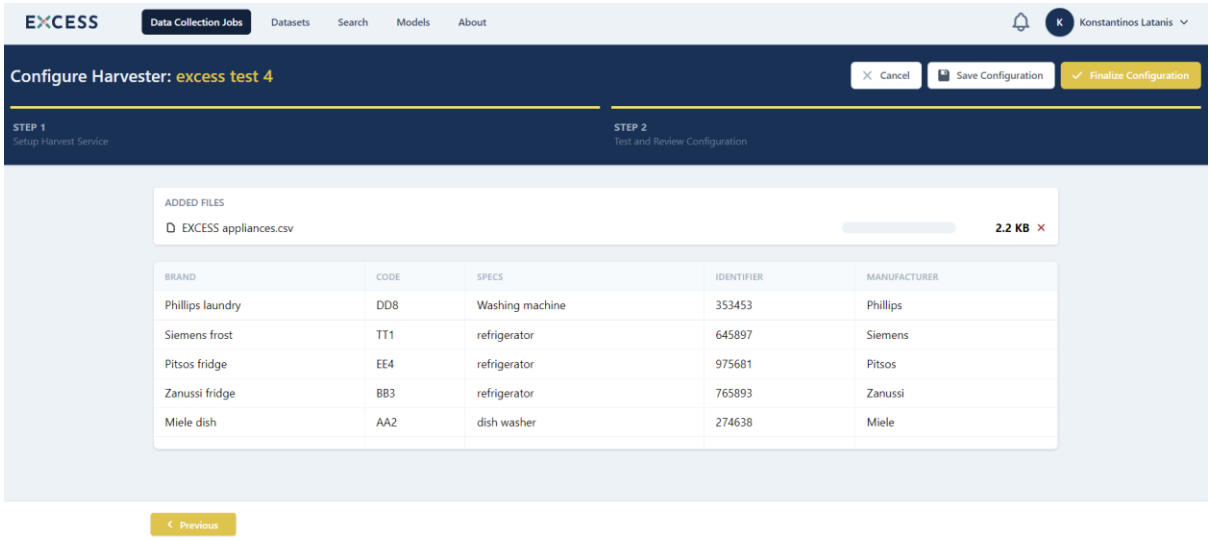

Figure 10-5: Check the data sample

The user has the option to save the data collection job, while a corresponding popup message is displayed, and revisit it later at the same stage.

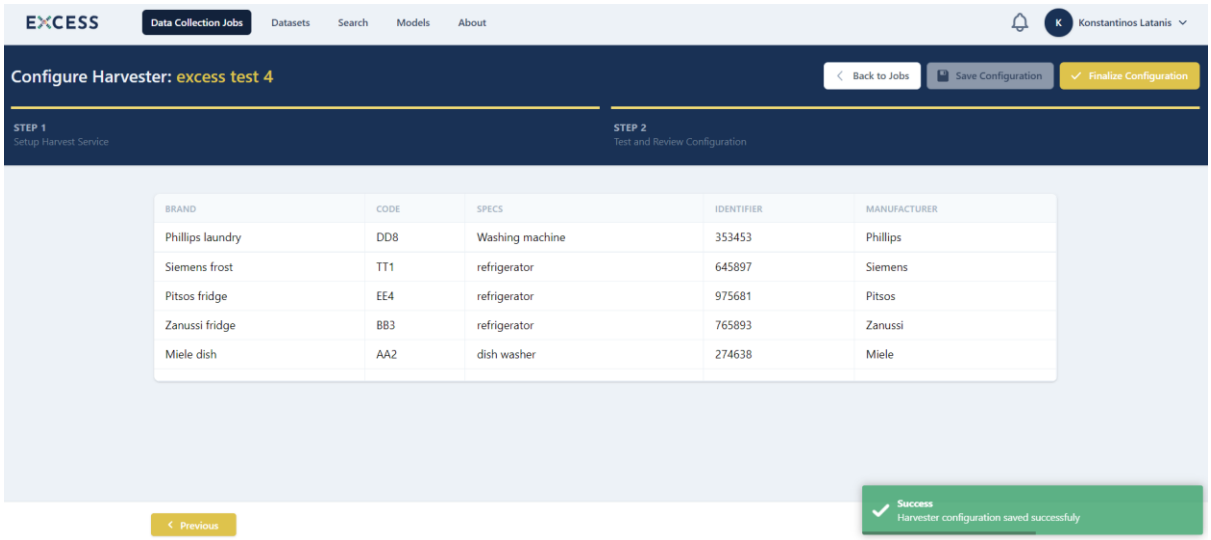

Figure 10-6: Save the data collection configuration

10.1.2 Acquisition through APIs

When the user chooses to configure an API for data acquisition, s/he is prompted to specify the settings of such an API connection. The user firstly selects the API response format between JSON and XML. The authentication details are defined, choosing among no authentication needed, bearer authentication and custom authentication.

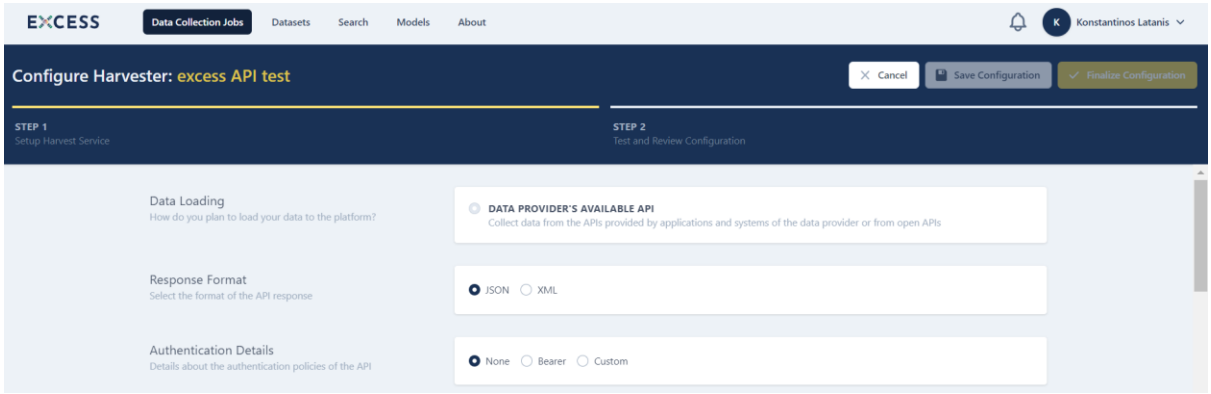

Figure 10-7: Configure API connection (1)

The user also defines the API URL and the method (GET, POST, PUT) along with the necessary request parameters as dictated by the API URL. In case of POST or PUT methods, the query body of the API connection has to be specified. In addition, the API pagination is defined among the options of no pagination, offset pagination and page pagination. For offset and page options, the related request parameters are also described.

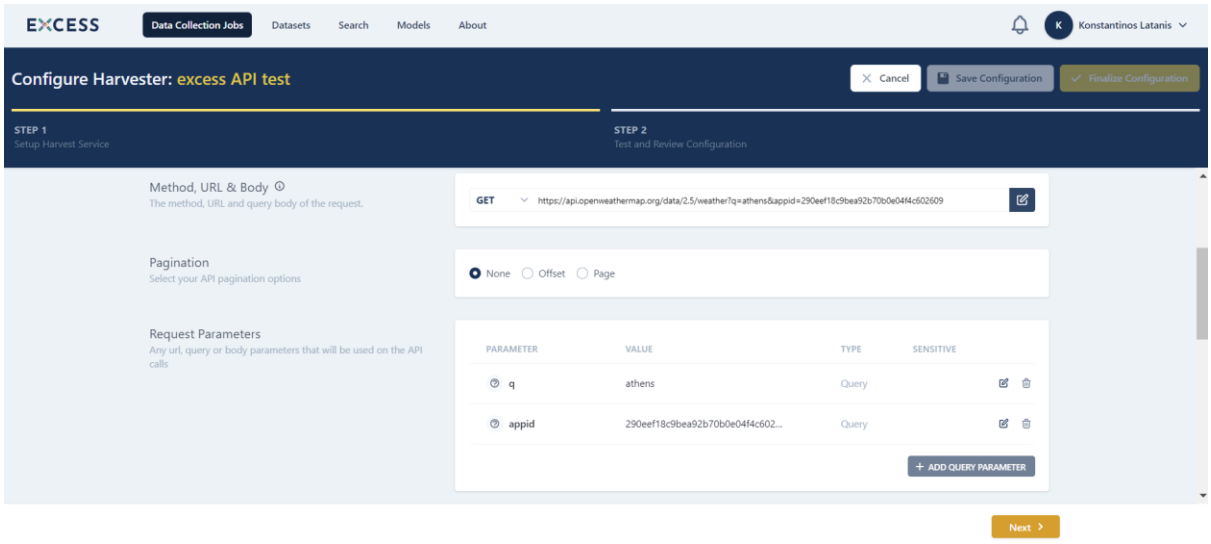

Figure 10-8: Configure API connection (2)

The user also describes the extra headers that may be needed for the realization of the API connection. Moreover, the retrieval settings are specified, as the user selects whether the data will be collected (a) once, by setting a specific date, (b) periodically, by setting a start and an end date along with a retrieval schedule depending on the selection of hourly, daily, weekly or monthly periodic collection or (c) every 60 seconds (polling), by setting a start and an end date.

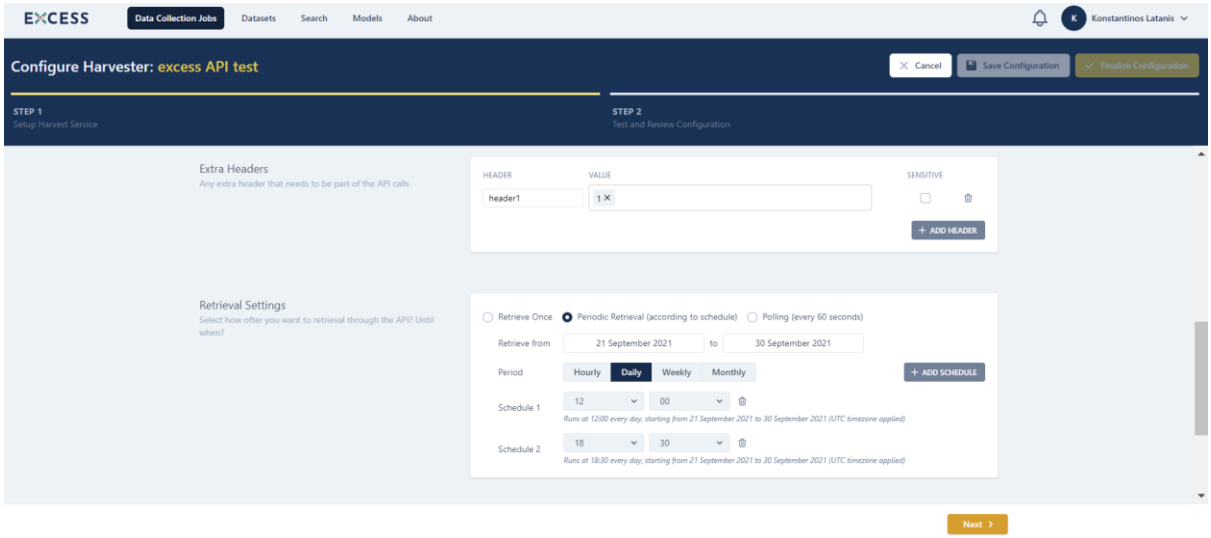

Figure 10-9: Configure API connection (3)

Moreover, the user selects how often the processing of the collected data will be carried out (immediately, every hour, every day).

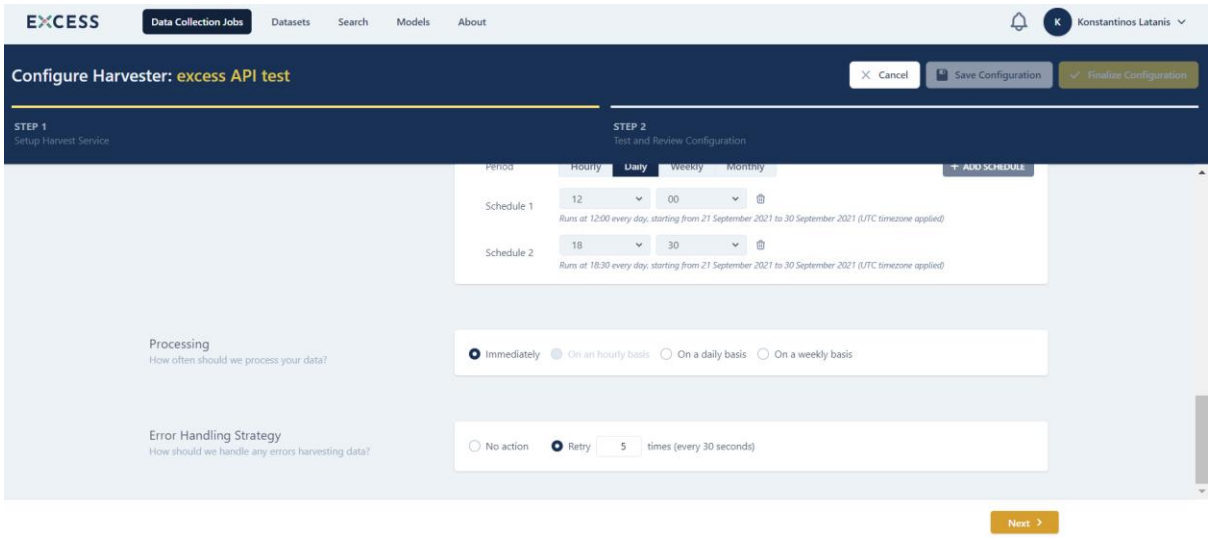

Figure 10-10: Configure API connection (4)

When the configuration of the API connection details is completed, the complete API response is presented according to the testing of the API connection. The user can choose whether the API response will be handled as a single record or as multiple records under a specific path and may add any additional static or dynamic parameters within the data. S/he may select which data elements of the API response will be further processed in the next stages of the EXCESS Data Management Platform and a preview of the selected actual data is displayed. The unselected data elements will be discarded and subsequently not be stored in the EXCESS Data Management Platform.

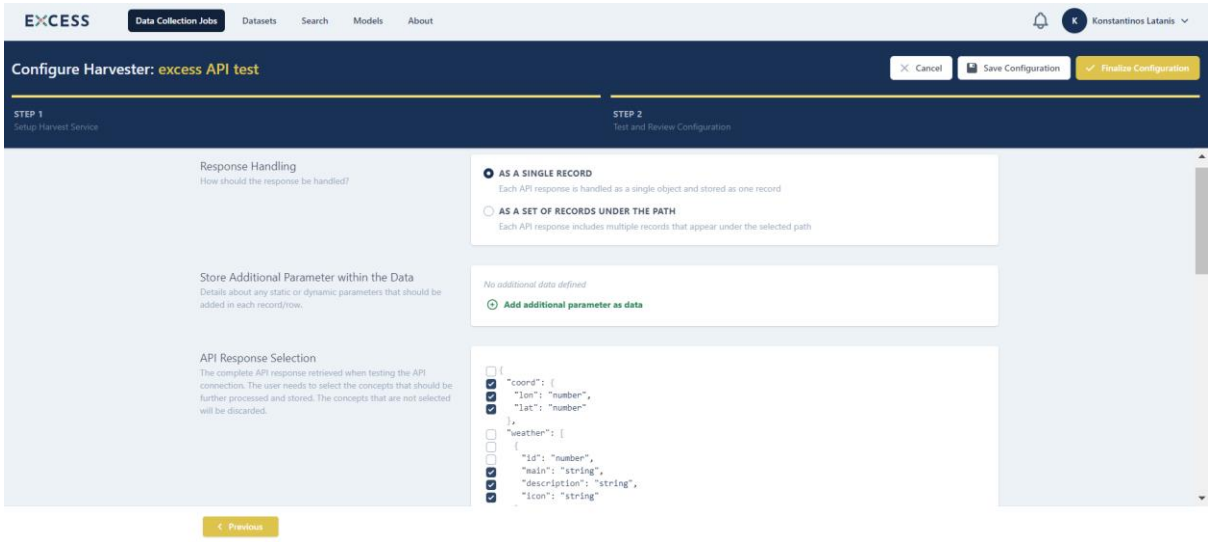

Figure 10-11: Test API connection and review API response

10.1.3 Acquisition through Pub/Sub mechanisms

When the user chooses the Pub/Sub mechanism for acquisition of streaming data, s/he firstly defines the format of the data to be streamed (JSON or XML). Moreover, the user specifies the connection URL, the Kafka topic, the ID of the group and the SASL mechanism along with any necessary credentials.

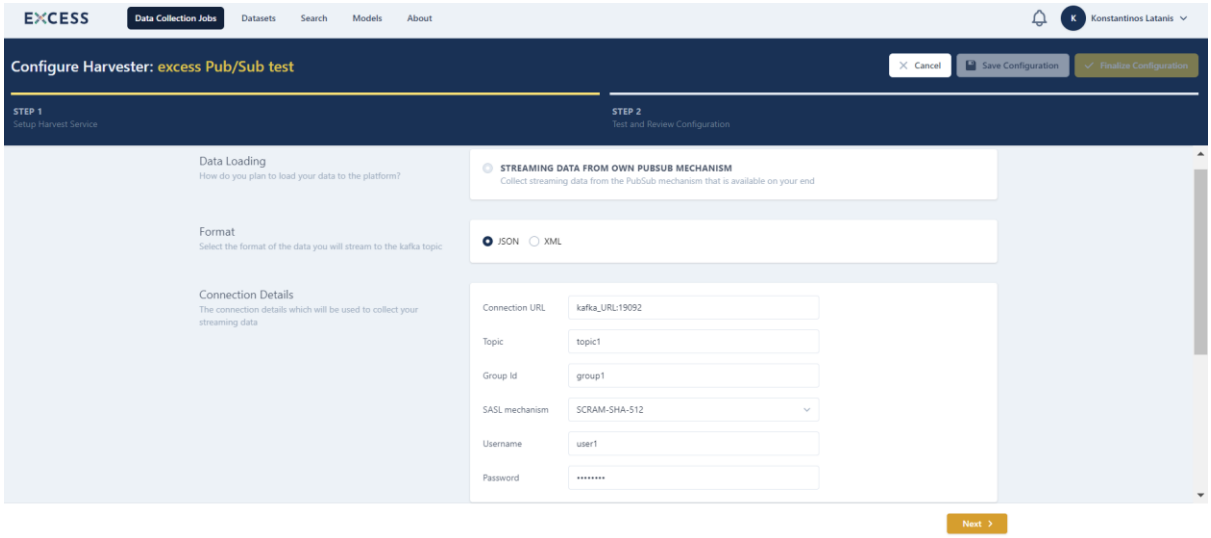

Figure 10-12: Configure Pub/Sub connection (1)

The user also defines until when the streaming data will be retrieved and how often they will be processed (every hour, every day).

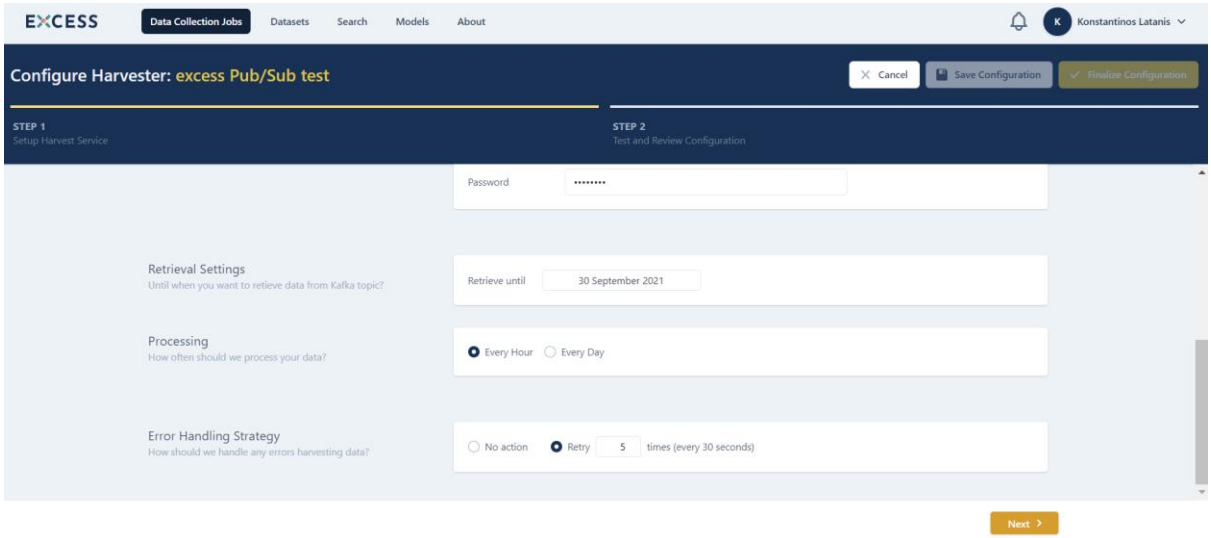

Figure 10-13: Configure Pub/Sub connection (2)

When the configuration of the Pub/Sub mechanism details is completed, the complete Pub/Sub response is presented according to the testing of the Pub/Sub mechanism. The user may select which data elements of the Pub/Sub response will be further processed in the next stages of the EXCESS Data Management Platform and a preview of the selected actual data is displayed. The unselected data elements will be discarded and subsequently not be stored in the EXCESS Data Management Platform.

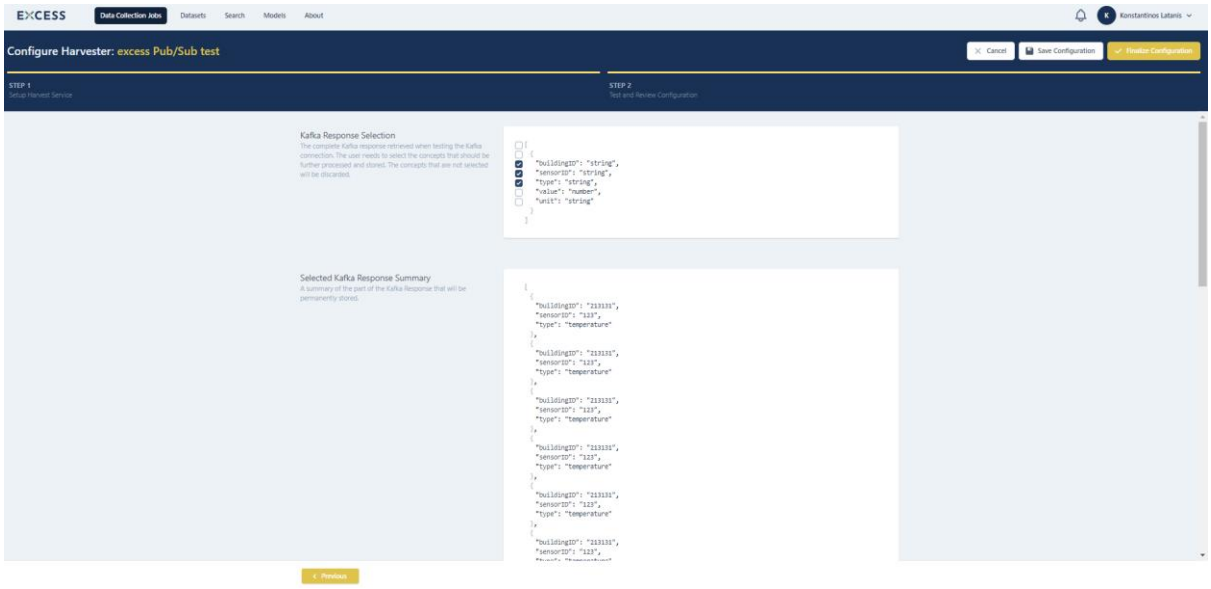

Figure 10-14: Test Pub/Sub connection and review Pub/Sub response

The user can save the Pub/Sub configuration in order to revisit it at a later time or finalize it. In case of finalization of Pub/Sub configuration, the user may change later only specific Pub/Sub mechanism details in order to maintain the Pub/Sub response consistency.

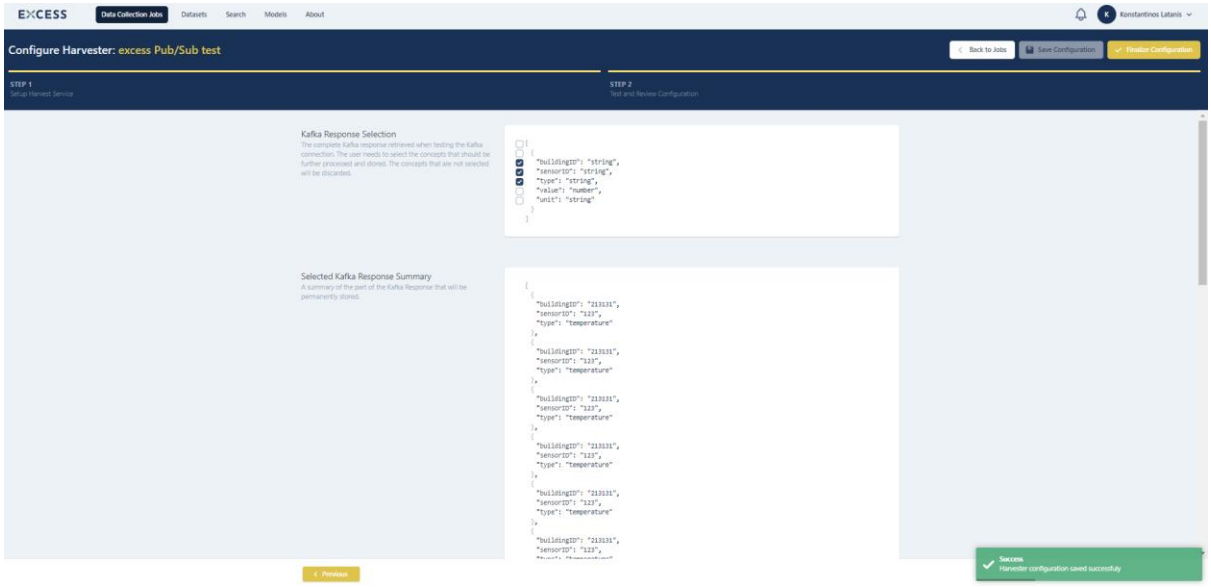

Figure 10-15: Check selected Pub/Sub response summary and save Pub/Sub configuration

10.2 Data Mapping

After the completion of the Data Collection step, the Data Mapping step follows, as long as the user has chosen it during the initial configuration of the pre-processing steps of the data collection job. If the user does not select the mapping procedure at that point, then the uploaded data will be stored in the EXCESS Data Management Platform as a single object and analysis activities that need data mapping will not be doable.

At the first step of the data mapping process, the user selects the domain where the data belong to and any standard that the data comply to, while s/he also chooses which category the data refer to. The domain actually comprises the EXCESS Common Information Model and the categories that the user may select from are the main concepts of the Model.

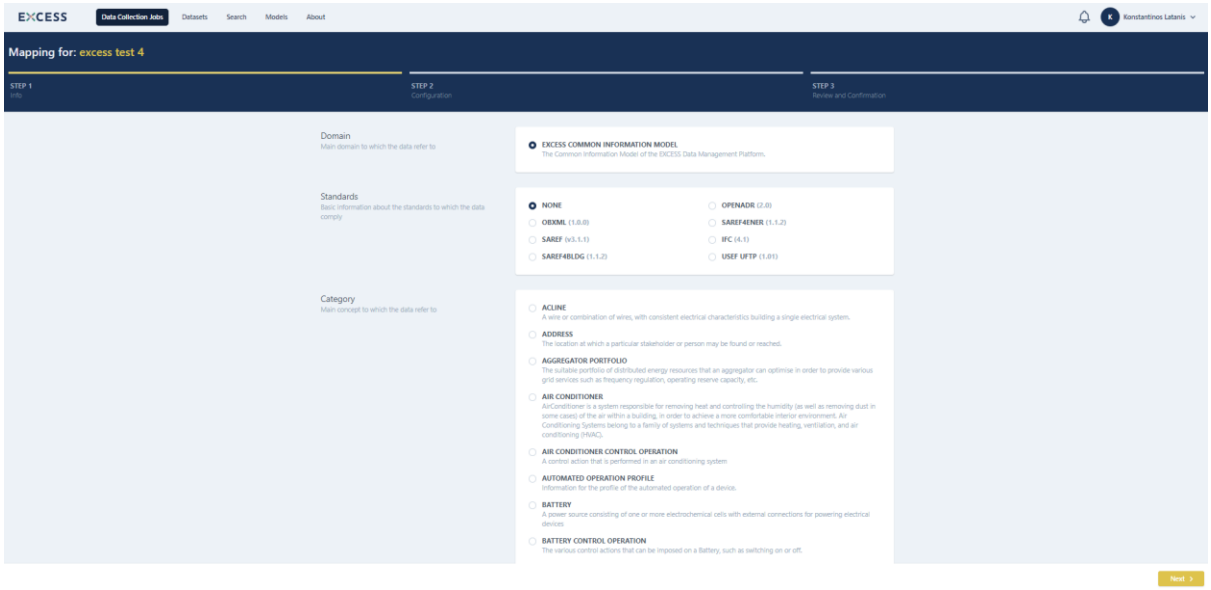

Figure 10-16: Select the EXCESS CIM, standard and category for mapping

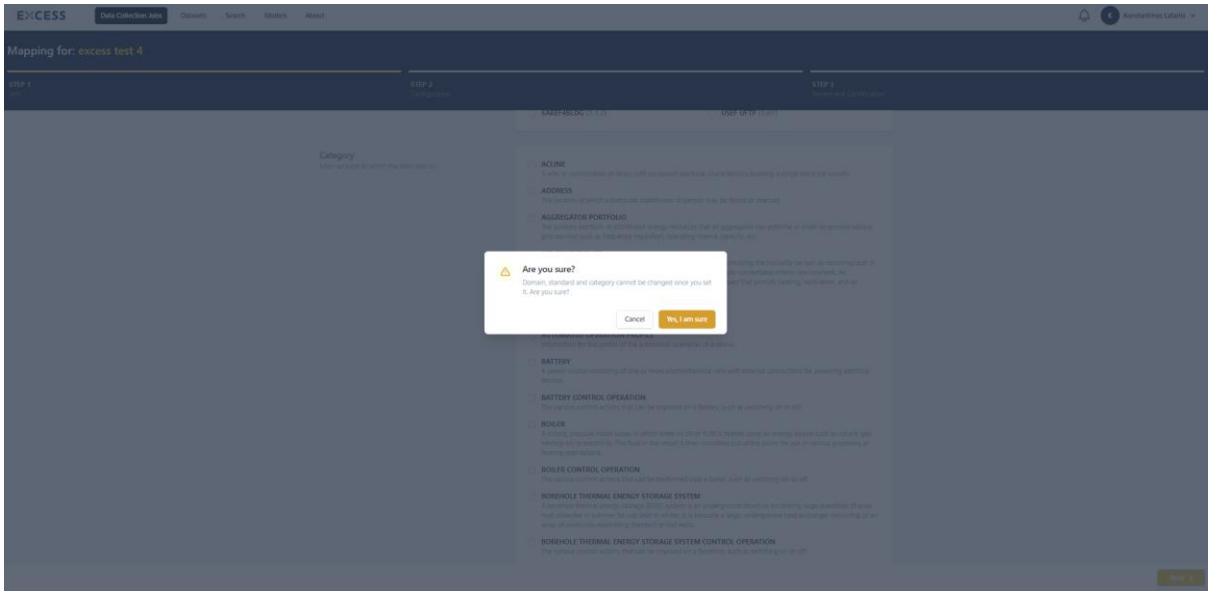

Figure 10-17: Confirmation of EXCESS CIM, standard and category selections

Upon finalizing and confirming the aforementioned selections, the semi-automatic mapping predictions of the EXCESS Data Management Platform are performed and the matchings of the data elements of the collected dataset to the equivalent concepts of the EXCESS Common Information Model are presented along with a level of confidence. The user may review the mapping results and identify any data type mismatches. The reviewing of mapping results is facilitated by the filtering of results to predicted, corrected, unidentified, invalid or selected. In the left pane in the Data Model area, the user can view the previously selected category of the EXCESS CIM with its various concepts and by clicking on each of them s/he can see its description and data type, something that facilitates the manual mapping process, where necessary.

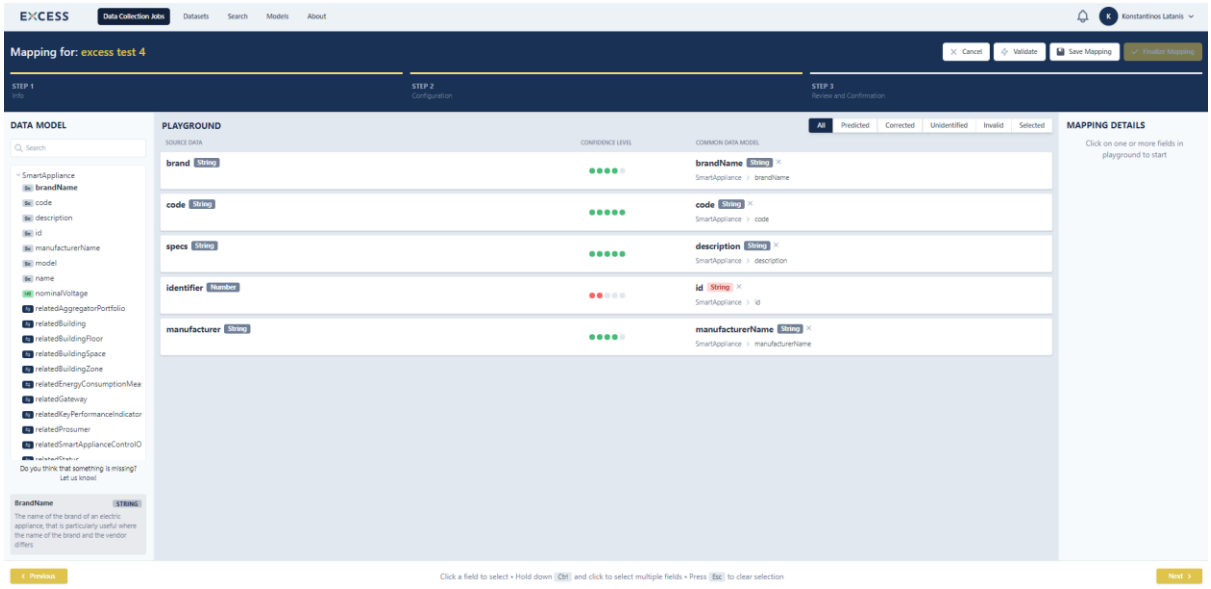

Figure 10-18: Review the mapping predictions

The user can view the details of a mapping in the right pane, by clicking on this mapping that is highlighted in blue, as well as its sample values under the Playground area. The user has to provide the measurement unit, in case a mapped concept has a numerical data type, so that the respective transformation to the baseline measurement unit of the EXCESS Common Information Model can be performed. In the same sense, in case a mapped concept has a datetime data type, the user has to specify the datetime format and the reference timezone of the data.

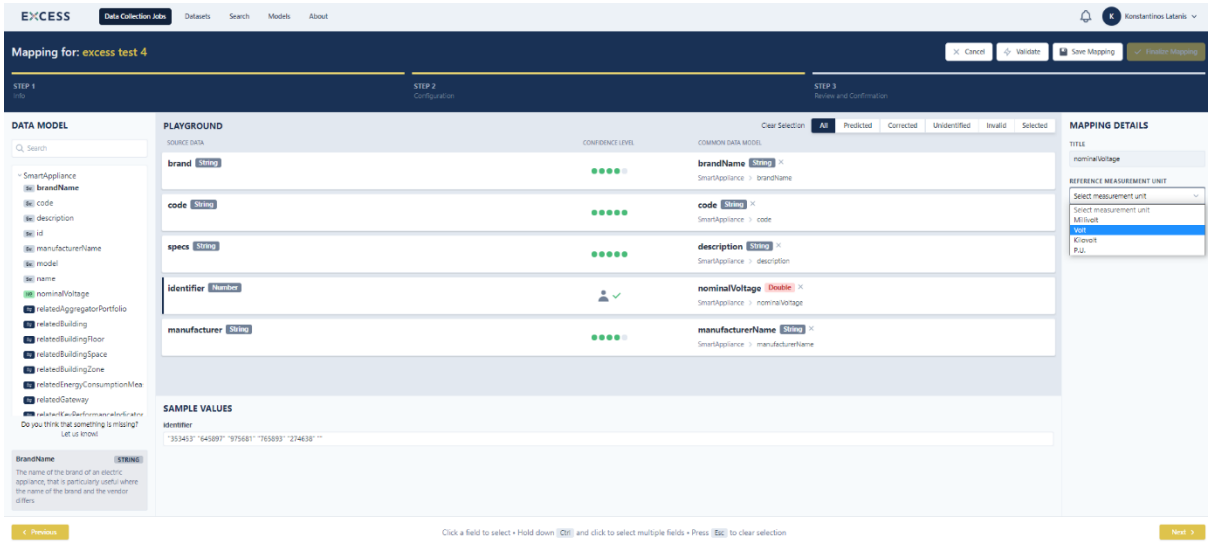

Figure 10-19: Define mapping details (1)

When the user selects one or more unidentified elements, s/he can also define a related concept pertaining to these elements. After choosing the desired related concept and providing the respective prefix, the user may set this related concept to continue searching for a manual mapping among more related concepts or request for a new prediction for the selected element(s) based on this specific related concept.

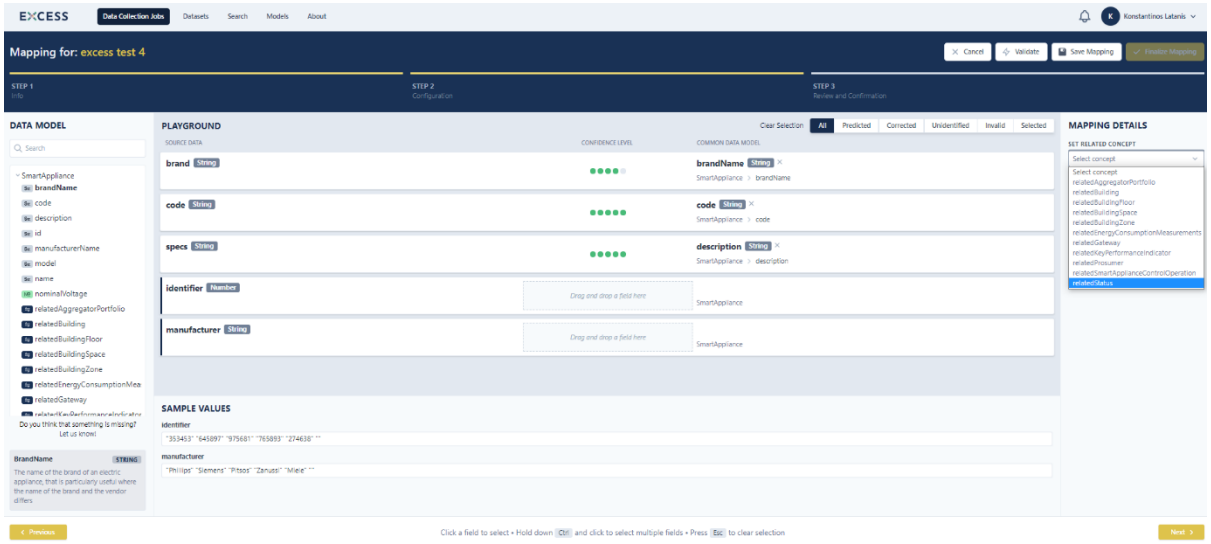

Figure 10-20: Define mapping details (2)

Once the new prediction is displayed for this element, the user may also see the related concept and its subconcepts in the Data Model area in the left pane.

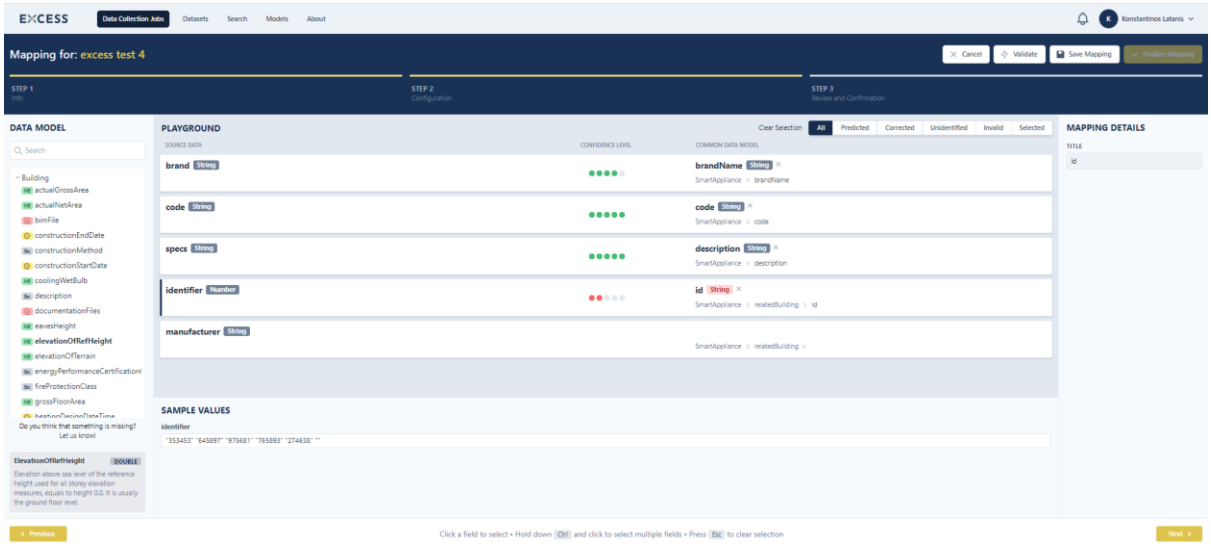

Figure 10-21: Navigate to related concepts and subconcepts

The user may also delete any mapping predictions to select manually concepts that match better to the data elements or may match any initially unidentified data elements to suitable concepts of the EXCESS Common Information Model. The manual mapping is performed by dragging and dropping a concept from the Data Model area to an unmapped data element in the Playground area. In case any unidentified elements are not mapped, they will be discarded from the rest of the steps of the data collection job and eventually will not be stored in the EXCESS Data Management Platform.

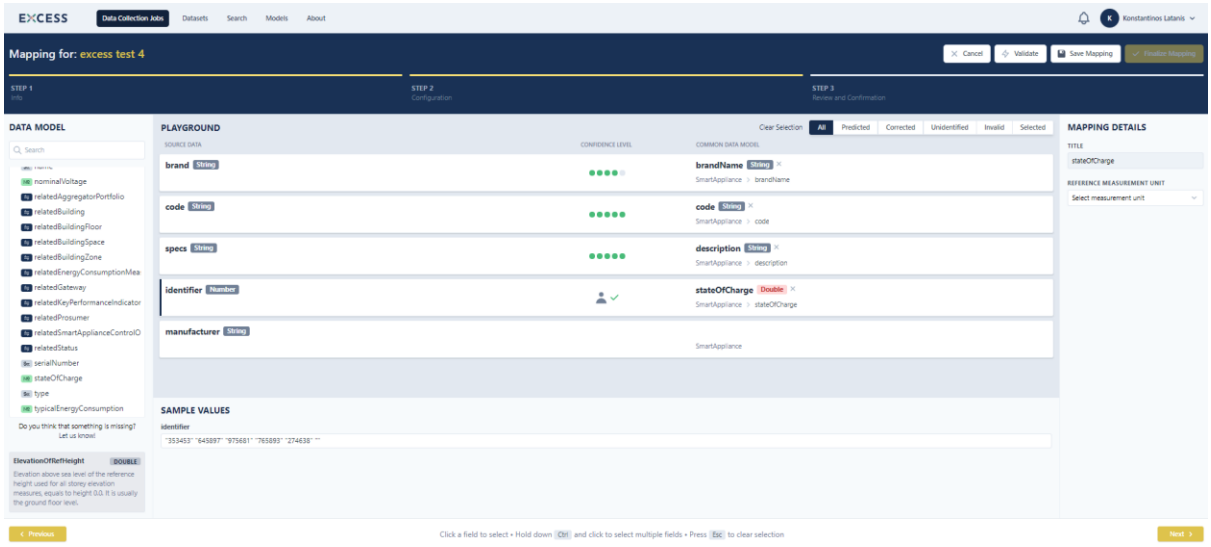

Figure 10-22: Define manual mappings

When the user proceeds to the final mapping stage, s/he can see an overview of the mapped and unidentified concepts and finalize the mapping configuration, which is saved in the configuration file. Upon finalization and confirmation, the mapping process cannot be edited anymore and the data collection job continues with the rest of its steps.

10.3 Data Cleaning

After the completion of the Data Mapping step, the Data Cleaning step follows, as long as the user has chosen it during the initial configuration of the pre-processing steps of the data collection job. At the first step of the data cleaning process, the user can view all the mapped concepts of the ingested dataset that have proceeded to the cleaning step. By clicking on any of them, the user can view their sample values, while s/he can make (or clear) multiple selections of these fields manually or by data type.

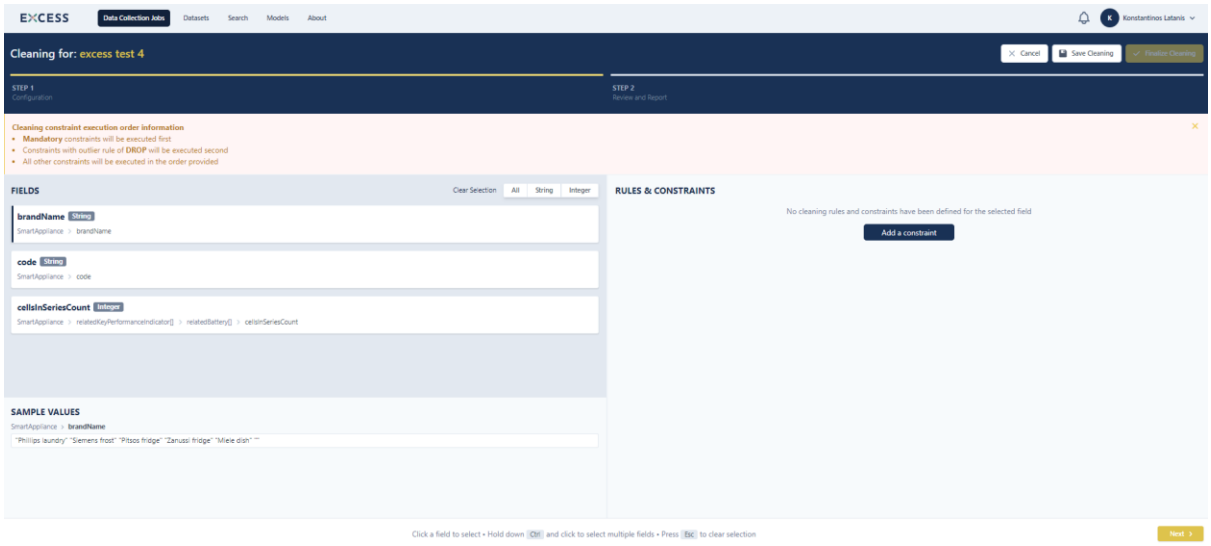

Figure 10-23: List of mapped concepts available for cleaning rules

By selecting one or more fields, the user is allowed to set the desired cleaning rules according to the data type of the selected fields. For each field, multiple cleaning rules may be applied. The user may also save the cleaning configurations to revisit at a later time.

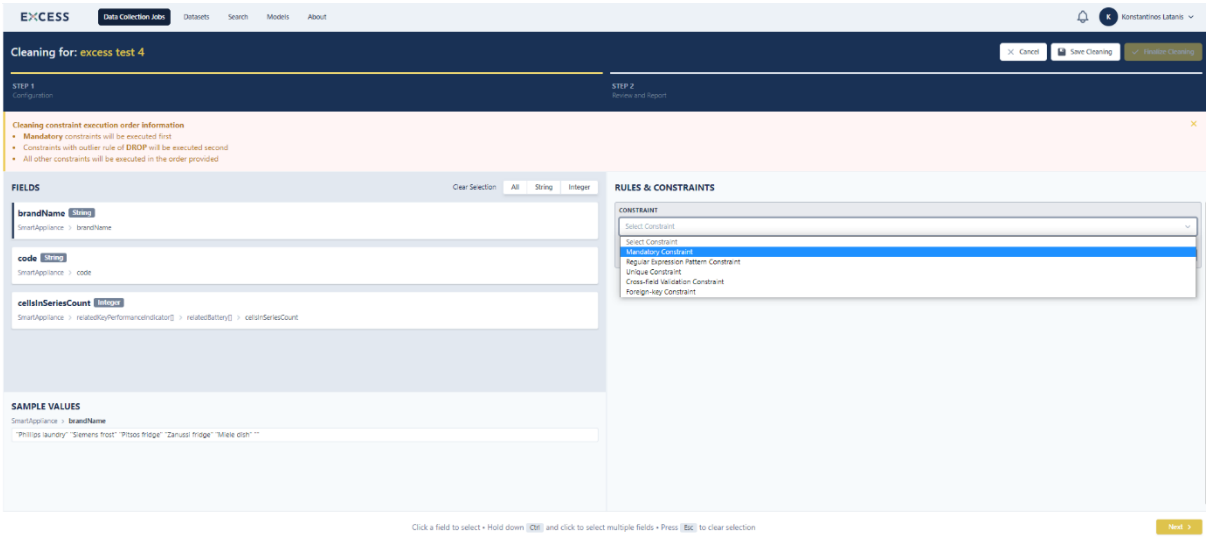

Figure 10-24: Set a cleaning rule (1)

For each cleaning rule, the user can specify its details and define how the outlier values will be handled by the Data Collection component. These outliers can be dropped from the dataset or replaced with other values selected by the user through a list of options offered depending on the data type of the field to be cleaned (e.g. replacement with a specific default value, the most frequent value of the field, the maximum value of the field, etc.). The user can edit or delete any created cleaning rule and s/he can also alter the execution order of cleaning rules for a specific field.

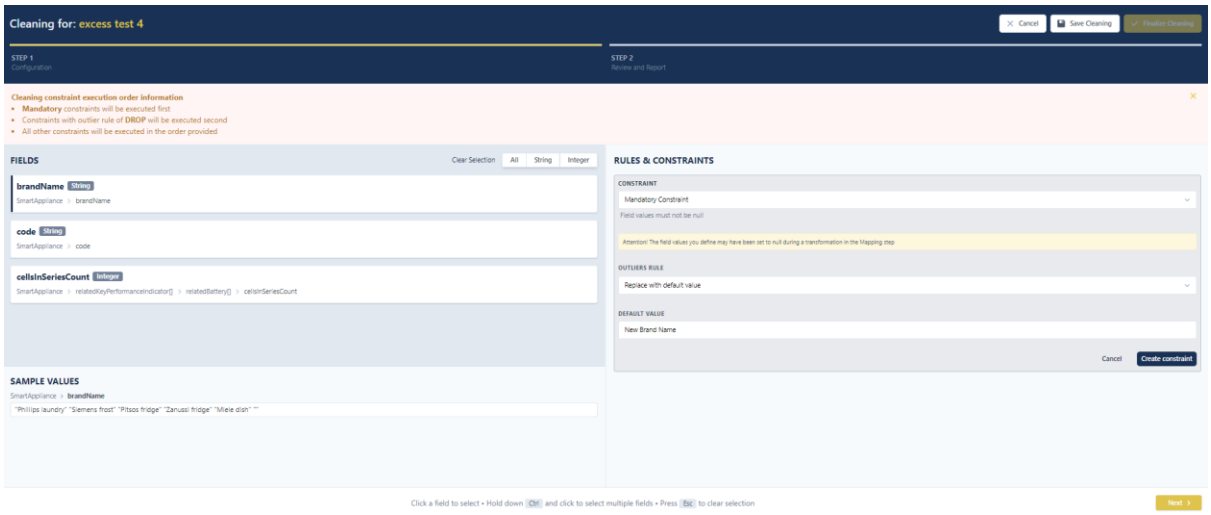

Figure 10-25: Set a cleaning rule (2)

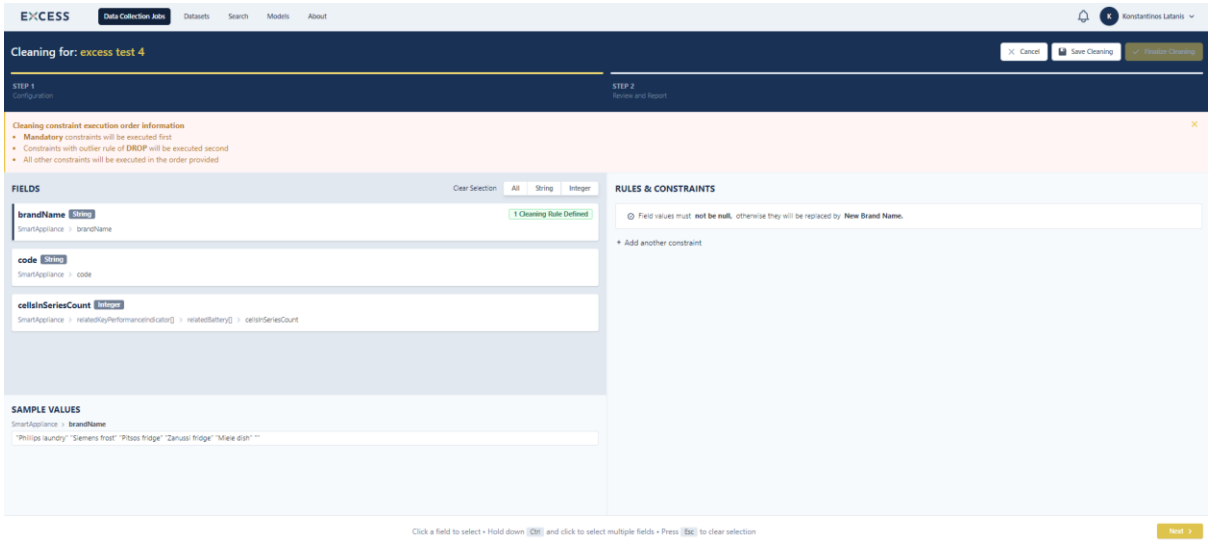

Figure 10-26: Set a cleaning rule (3)

As soon as the user proceeds to the final cleaning stage, s/he can see an overview of the assigned cleaning rules as well as the fields that will not undergo the cleaning process and finalize the cleaning configuration, which is saved in the configuration file. During the data cleaning process, the execution order of the assigned cleaning rules is the following: (a) Mandatory constraints are executed first, (b) Constraints with "drop" outlier rule are executed second, (c) All other constraints are executed in the order provided by the user. Upon finalization and confirmation, the cleaning process cannot be edited anymore and the data collection job continues with the rest of its steps.

10.4 Data Anonymization

After the completion of the Data Mapping or Data Cleaning step, the Data Anonymization step follows, as long as the user has chosen it during the initial configuration of the pre-processing steps of the data collection job. At the first step of the data anonymization process, the user can view all the mapped concepts of the ingested dataset that have proceeded to the anonymization step, while by clicking on any of them, the user can view their sample values. By default, all fields are initially insensitive.

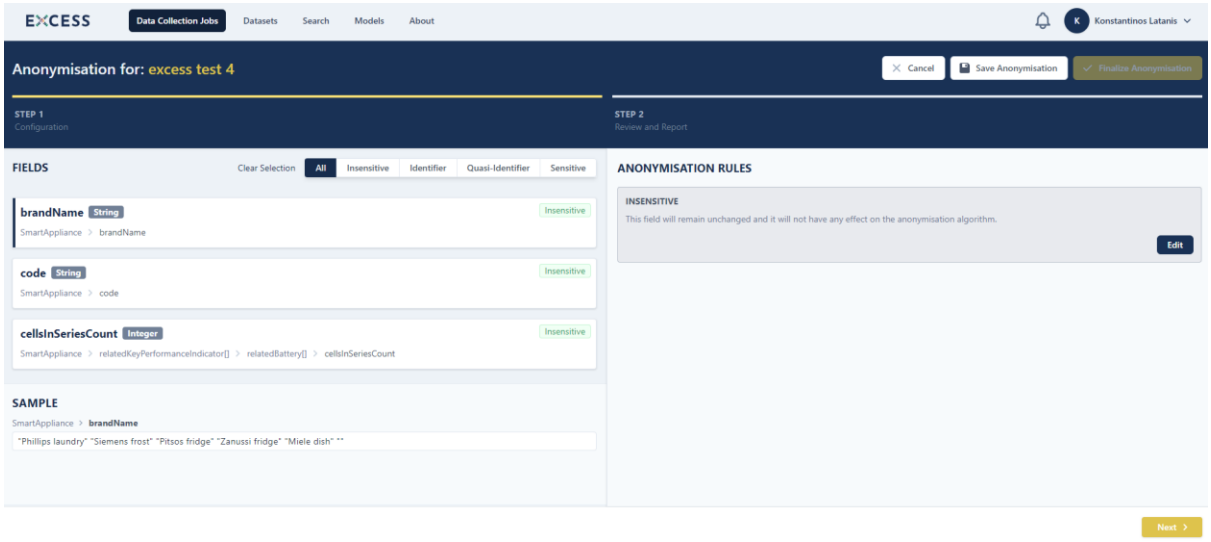

Figure 10-27: List of mapped concepts available for anonymization rules

By selecting a field, the user is enabled to set an anonymization rule with its details depending on the data type of the field. The user may (a) maintain a field as insensitive, thus no anonymization rule will be applied, (b) set the field as identifier, where the respective column will be dropped, (c) set the field as quasi-identifier, where the appropriate generalization method will be applied according to the field data type (masking for string data type and interval or numerical group for numerical data type) and (d) set the field as sensitive, so that the anonymization algorithm will protect these field values by generalizing the quasi-identifiers. The user has also the option to filter the fields based on the aforementioned four categories of assigned anonymization rules. Especially for the generalization method of quasi-identifiers, the user is enabled to see an example of the generalization levels that may be applied based on the anonymization algorithm. The user can edit any of the created anonymization rules and also save the anonymization configurations to revisit at a later time.

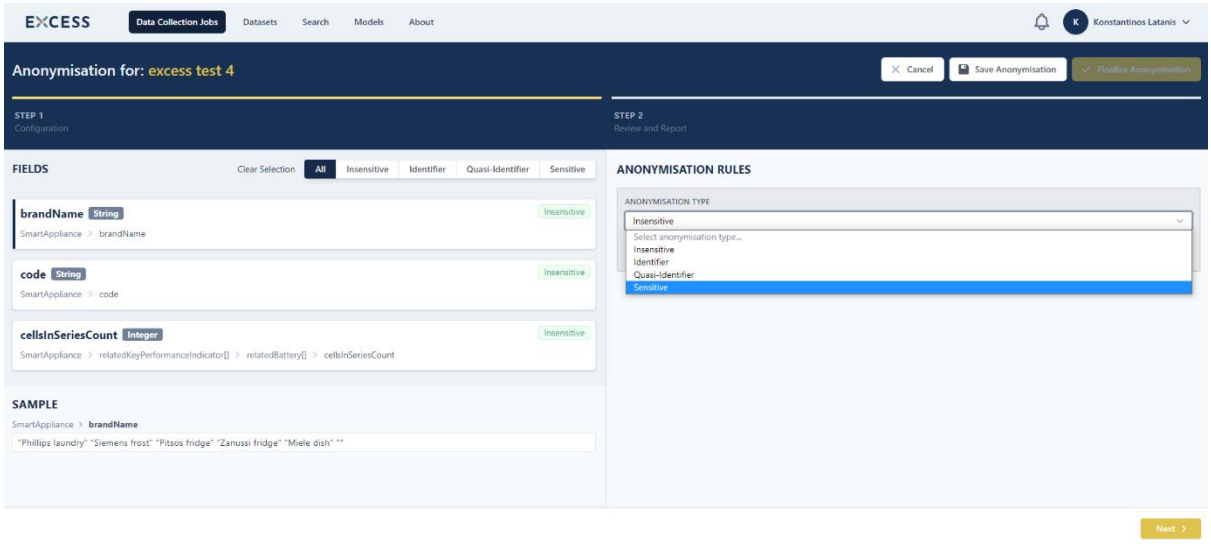

Figure 10-28: Set an anonymization rule (1)

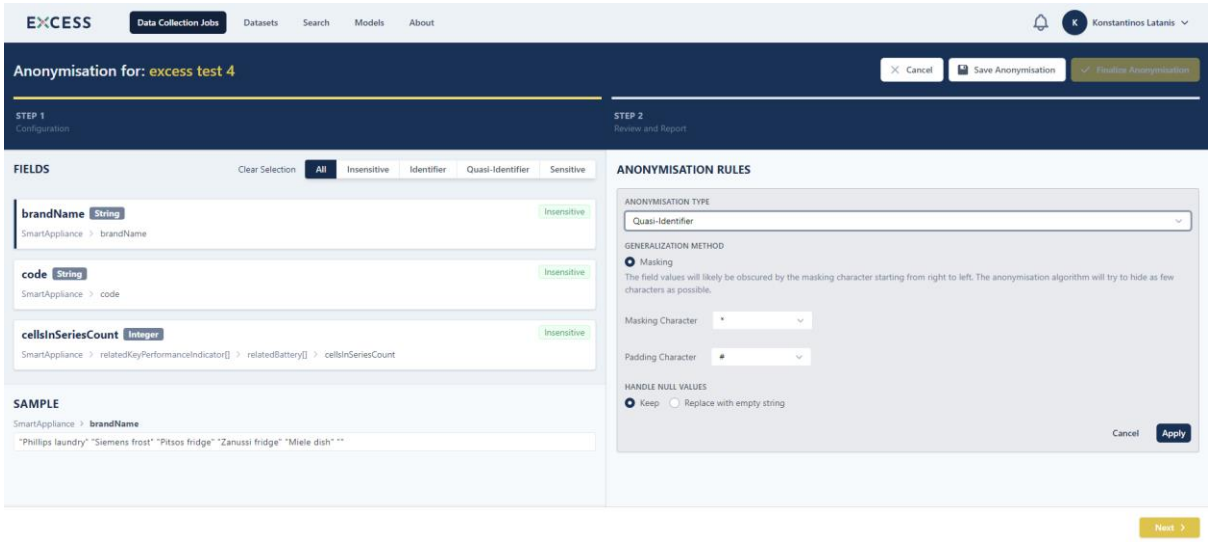

Figure 10-29: Set an anonymization rule (2)

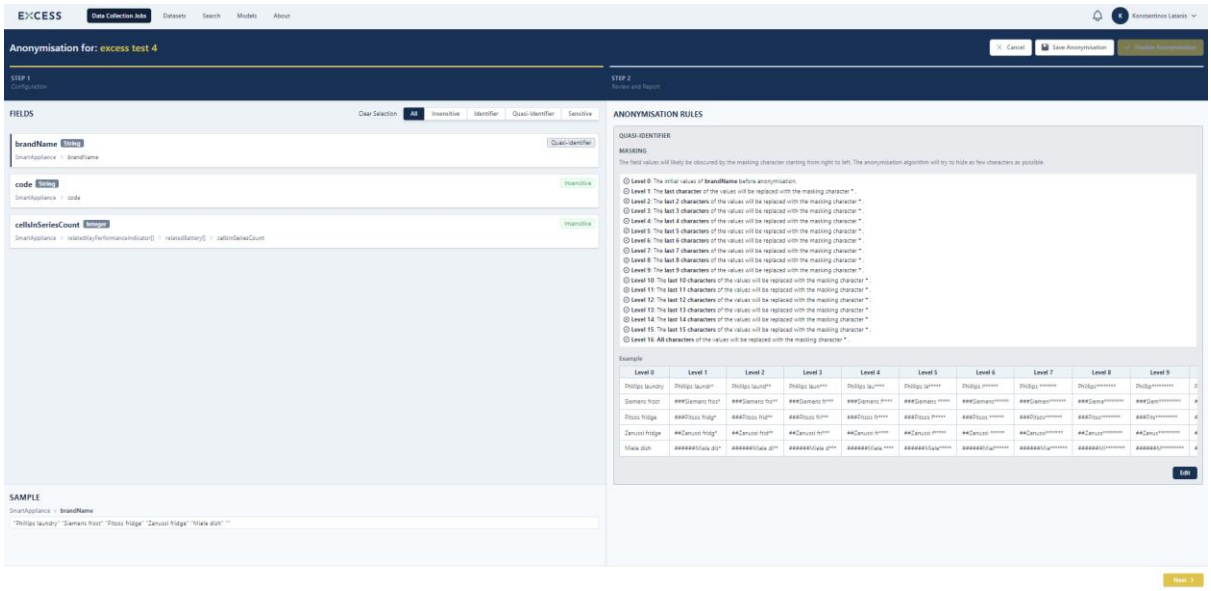

Figure 10-30: Set an anonymization rule (3)

As soon as the user proceeds to the final anonymization stage, s/he can see an overview of the anonymization rules as well as the fields that will remain insensitive and finalize the anonymization configuration, which is saved in the configuration file. The user may provide the acceptable information loss threshold that if it is exceeded, the anonymization process will fail since the data will be destroyed upon the designated threshold. Moreover, s/he can define the level of k-anonymity, which is the data anonymization algorithm. Upon finalization, the anonymization process cannot be edited anymore and the data collection job continues with the storage process.

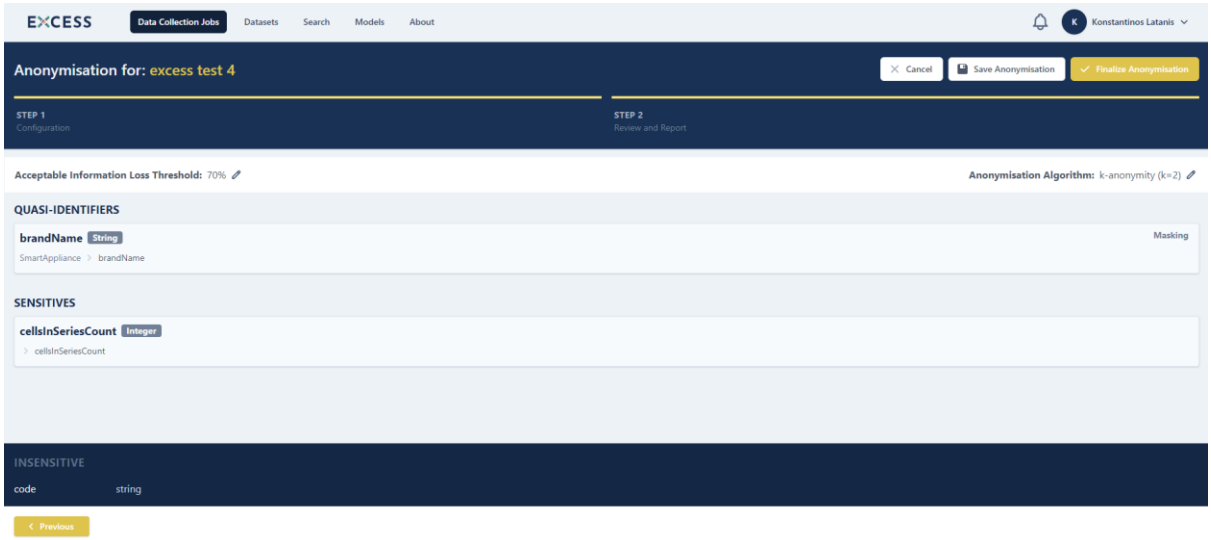

Figure 10-31: Summary of assigned anonymization rules

10.5 Data Storage and Update

As long as a dataset has undergone all the designated by the user pre-processing steps, it arrives at the Data Storage step, where the data are finally loaded in the EXCESS repositories.

The user is prompted to define the name and the description of the dataset and upon finalization, the data are finally stored in the repositories of the EXCESS Data Management Platform, while the data collection job is completed.

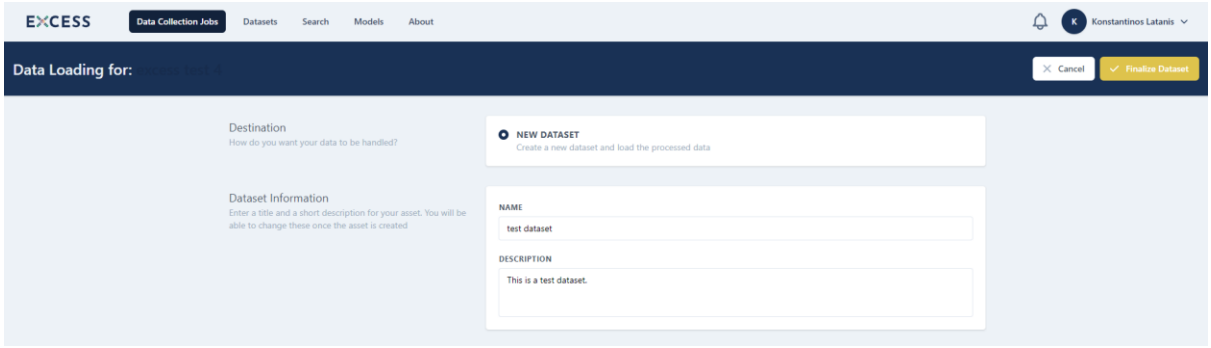

Figure 10-32: Define name and description of dataset

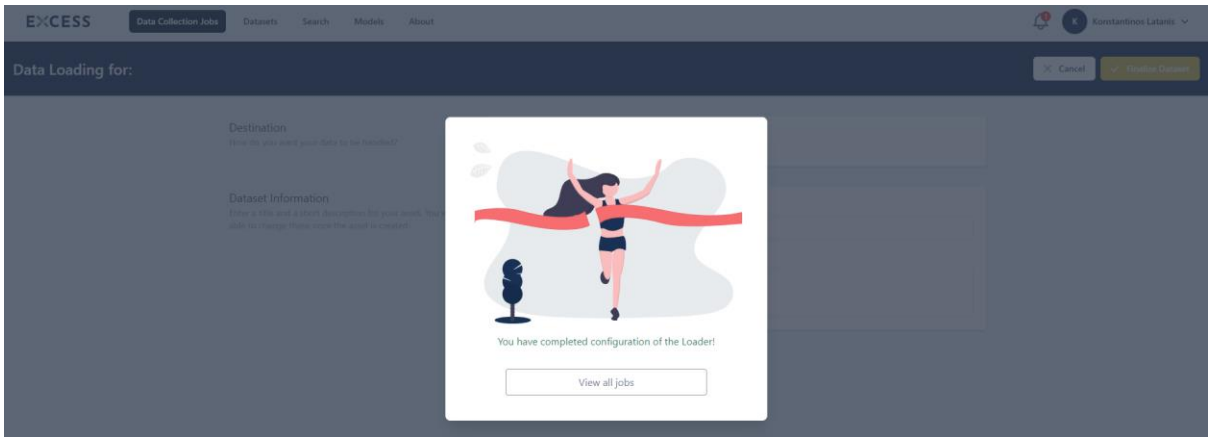

Figure 10-33: Completion of data collection job and data stored in EXCESS Data Management Platform

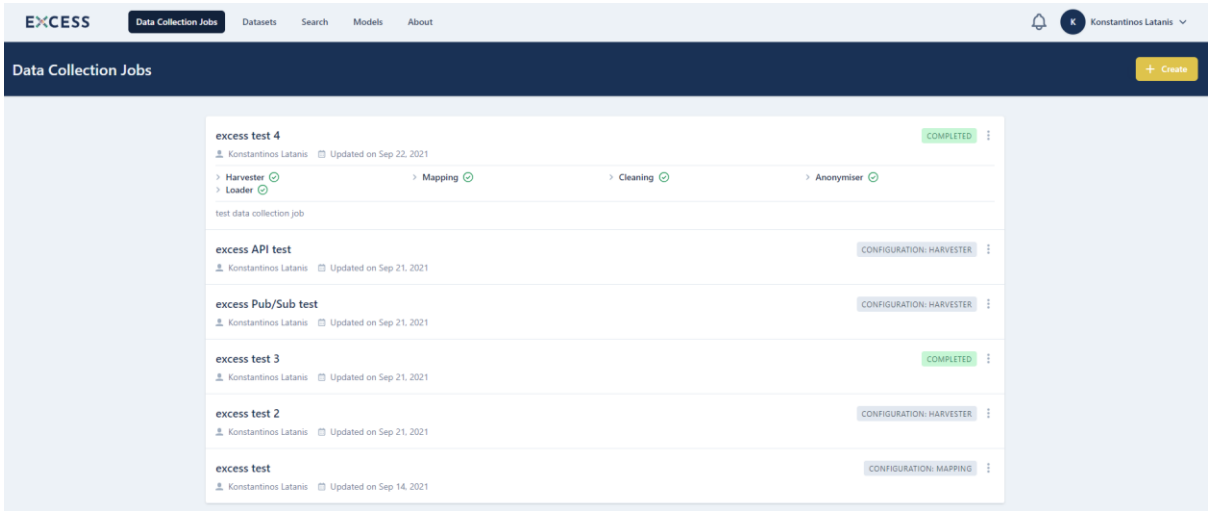

Figure 10-34: Data collection job completed

In the case of file uploading, the user may update a dataset by choosing its previously executed data collection job and adding a new file with the same data structure in the data collection step. The already created configuration file for this job with all its previously defined pre-processing rules will be used and the dataset will be finally appended with the data of the new file. The data update is not feasible only in case specific cleaning rules had been assigned in the previously executed data collection job or the initial dataset had undergone the anonymization process.

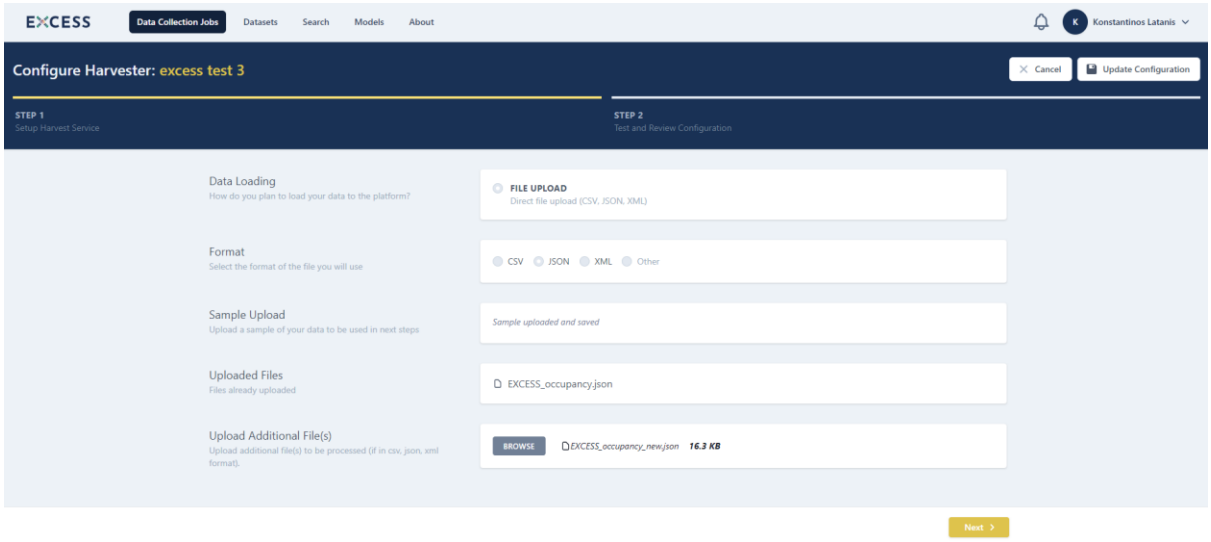

Figure 10-35: Update a dataset (1)

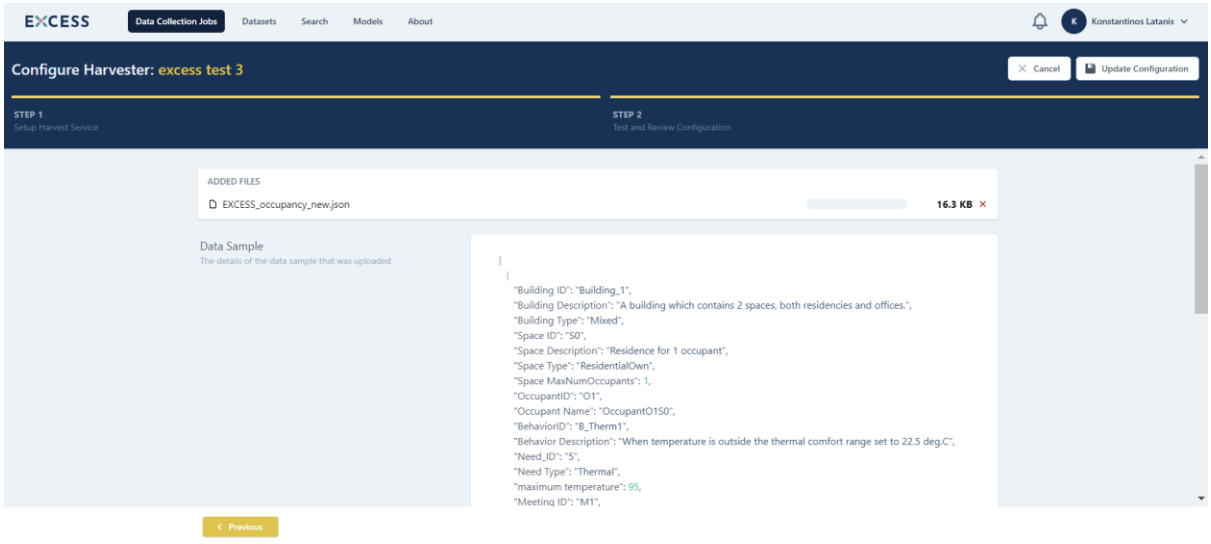

Figure 10-36: Update a dataset (2)

10.6 Metadata definition

When a data collection job is completed, the stored dataset is displayed in the list of datasets as incomplete because its metadata are needed to be defined by the user in order to make it available to the EXCESS Data Management Platform.

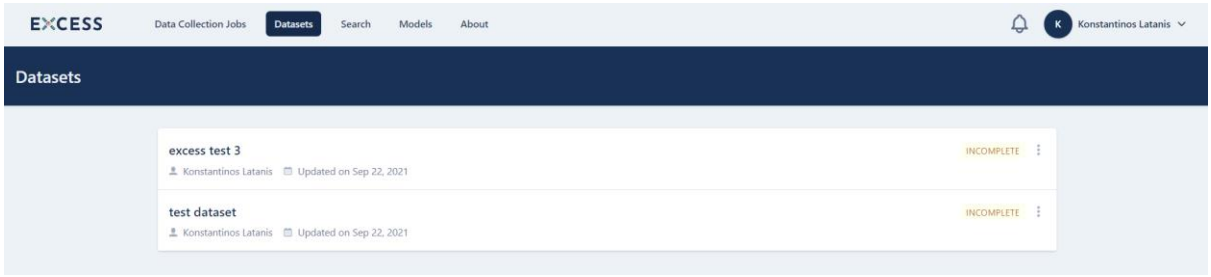

Figure 10-37: List of datasets

By clicking on this dataset, the user may see its overview and data structure, while s/he is prompted to fill in its metadata.

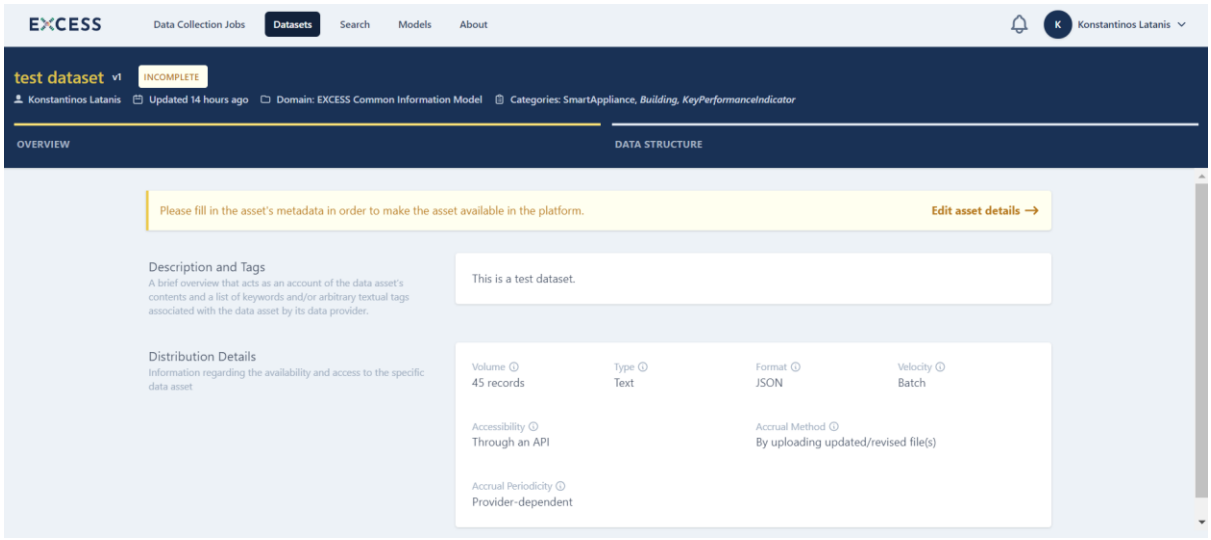

Figure 10-38: Overview of dataset

By clicking on the edit option, the user may define the general information of the dataset and its distribution, temporal and spatial details. Moreover, the licensing and access policy information of the dataset are specified, indicating which users can have access on this dataset.

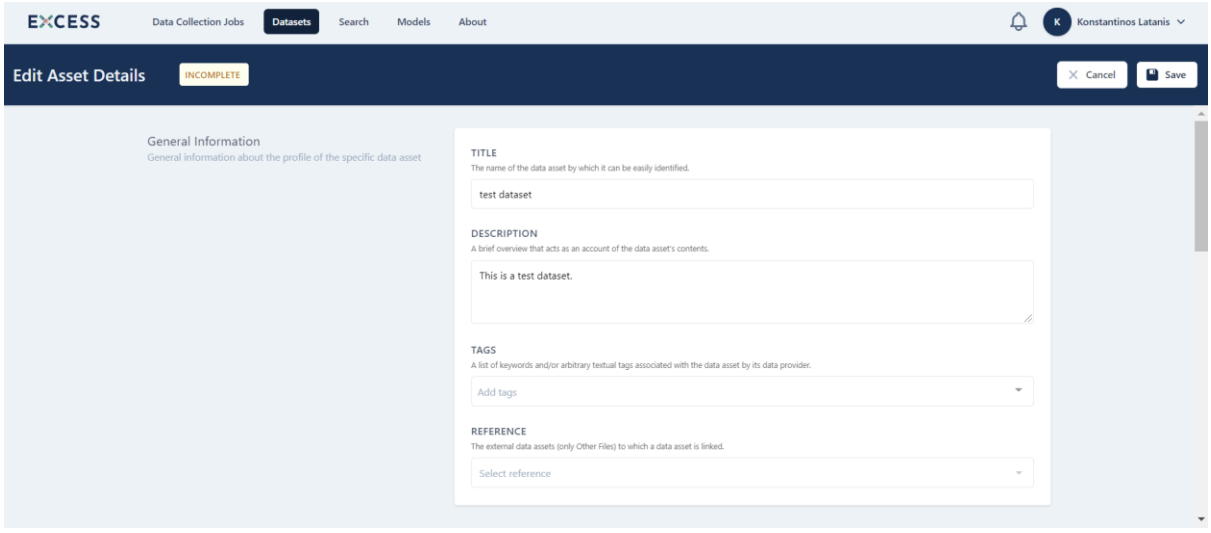

Figure 10-39: Edit dataset metadata (1)

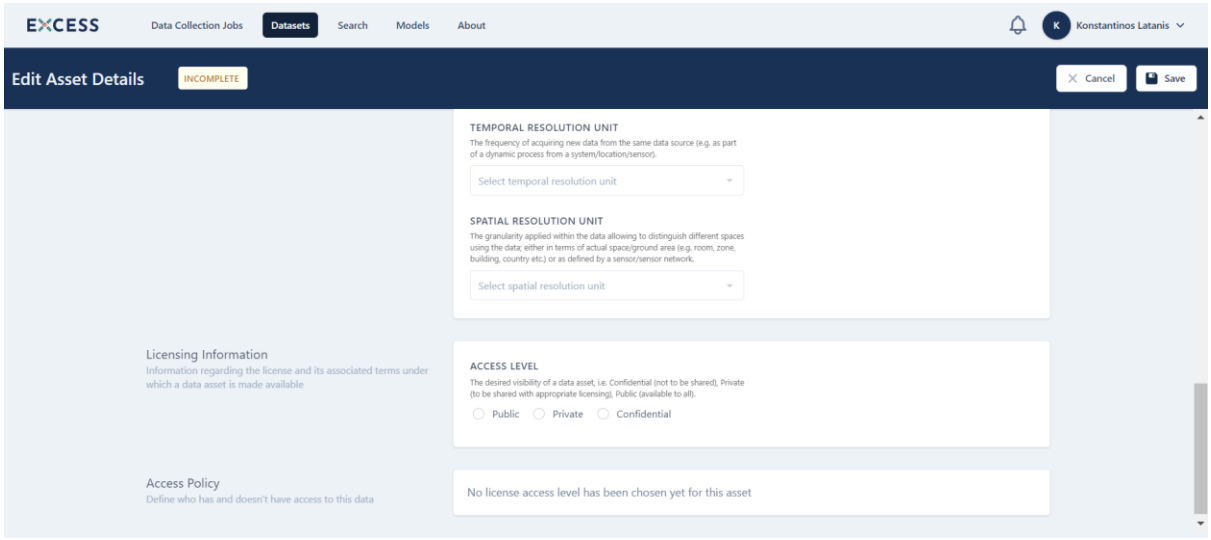

Figure 10-40: Edit dataset metadata (2)

Upon saving the metadata selections, the dataset is displayed as available in the list of datasets and can be used for analysis and other purposes in the EXCESS system.

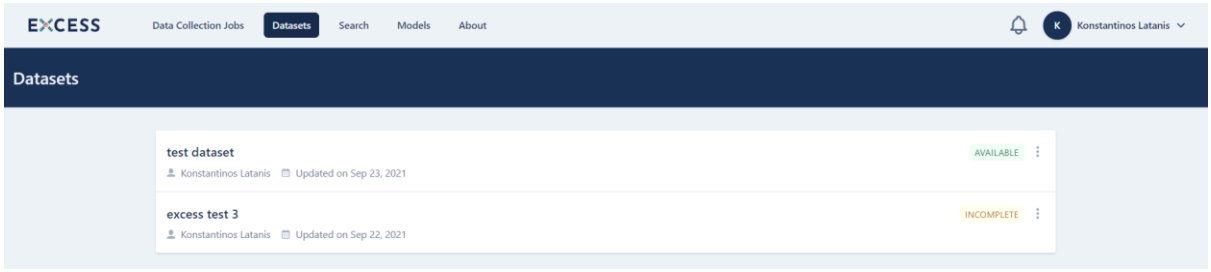

Figure 10-41: Dataset available with completed metadata

11 Conclusions

This deliverable has documented the activities of the Task 3.2 “Interoperable Data Management Framework” towards the design and development of the EXCESS Data Management Platform and the elaboration of the EXCESS Common Information Model. The EXCESS Common Information Model comprises the common language for all the datasets residing in the EXCESS Data Management Platform so that they are eligible for analysis and other purposes. The methodology for the elaboration of the EXCESS Common Information Model has been presented, regarding the provision of sample datasets by the demo site partners, while the related building and energy standards that have been used as a basis for the model elaboration have been described. In this sense, the concepts of the EXCESS Common Information Model that is encapsulated in the EXCESS Data Management Platform have been provided. Moreover, the different components that comprise the EXCESS Data Management Platform have been described, namely the Data Collection component, the Data Mapping component, the Data Cleaning component, the Data Anonymization component, the Data Storage component and the User Management service. The functionalities and the related technologies that have been used for their implementation have been provided. In addition, a comprehensive guide for the use of the different components of the EXCESS Data Management Platform is presented, showing the corresponding collection, pre-processing and storage tasks.

The deliverable D3.2 has presented the first release of the EXCESS Data Management Platform and the EXCESS Common Information Model. In M32 of the project, an updated version of the deliverable will be elaborated, including all the refinements and updated functionalities of the final release of the EXCESS Data Management Platform along with the updated EXCESS Common Information Model based also on the feedback collected by the demonstrators during the initial operation of the buildings of the demo sites.

References

[1] ETSI (n.d.) New version of the Machine 2 Machine Standard for Smart Appliances introduced by the ETSI," Digital Single Market, Available online: https://saref.etsi.org/

[2] Daniele L., den Hartog F., Roes J., "Created in Close Interaction with the Industry: The Smart Appliances REFerence (SAREF) Ontology," (2015) in Formal Ontologies Meet Industry, Springer, Cham, pp. 100–112

[3] ETSI TS 103 410-1 V1.1.2 (2020) SmartM2M; Extension to SAREF; Part 1: Energy Domain. Available online https://saref.etsi.org/extensions.html#SAREF4ENER (accessed 19/05/2021).

[4] OPENADR Standard Specifications, Available online: https://www.openadr.org/

[5] USEF Energy – Universal Smart Energy Framework, Available online: https://www.usef.energy/

[6] IFC data model specifications, Available online: https://www.buildingsmart.org/standards/bsi-standards/industry-foundation-classes/

[7] [ETSI TS 103 410-3 V1.1.2 (2020) SmartM2M; Extension to SAREF; Part 3: Building Domain, Available online, https://saref.etsi.org/saref4bldg/v1.1.2/

[8] Green Building XML schema, Available online: https://www.gbxml.org/About_GreenBuildingXML_gbXML

[9] obXML Schema Specifications, Available online: https://behavior.lbl.gov/?q=obXML

Annex: Data Collection Exercise

Finnish demo site:

Data Description		Data Description		Data Asset Features														Data Asset Availability							Data Asset Rights		Data Analysis Types of analysis currently conducted on data	Data Asset Assessment Relevance to specific demonstration activities					Comments
Dataset ID	Data Asset Title	Description	Date of Row Last Update	Volume	Variety	Type	Format	Velocity	Historical Data Availability	Temporal Coverage	Spatial Coverage	Language	Relevant Standards	Veracity	Temporal Resolution	Spatial Resolution	Dependency / Linking to Other Sources	Data Asset Owner	Data Asset Available from 3rd Party	Data Asset Provider	Accessibility Method	Frequency of Updates	Update Strategy	Documentation	Privacy	Need for Anonymization		Accuracy	Completeness	Timeliness			
[Unique identifier following the convention 'Country_Partner_DA_#nr']	The title of the data asset	A brief description of the data asset - At least 2-3 lines to give an overview of the data	The date of the last update of this row in the specific spreadsheet	[X GBs / records / transactions per hour / day / month]	[Structured / Unstructured / Semi-structured]	[Text / Image / Other]	[csv, xml, json, other]	[Real-time, Near Real-time, Batch]	[Y/N]	[From ... To...]	[Locations]	[e.g. Spanish, Finnish, French, Dutch, German, English...]	[List the international standards to which a data asset complies]	[Raw, Pre-processed, Processed Data asset]	[<the temporal "granularity" of the data, e.g. per minute / hour / day / month]	[<the spatial "granularity" of the data, e.g. at district / zone / building / area level]	[Y/N, If Y, list the other sources or codelists]	The name of the data asset owner	[Y/N]	The name of the data asset provider in EXCESS	[Through API, As downloadable file, As database extract, Other]	[Real-time, Every X minutes / hours, Daily, Weekly, Monthly, Yearly, other]	[Append new data / Replace existing data / other]	The documentation of the API or data sample [incl. the location and the name of the file in the EXCESS repository]	[<considerations (no IS can be shared at all) / Proprietary (to be shared with appropriate licensing with the demonstrator partners) / Private (to be shared with appropriate	[Y/N depending on whether the data asset contains sensitive or personal data]	[e.g. Correlation analysis for..., Predictive analytics for...]	[measure of correctness and precision, e.g. whether the dataset is effo-free. Ranked 1 (Low) - 5	[Degree to which a data asset is sufficient in scope, depth. Ranked 1 (Low) - 5 (High)]	[How long a data asset remains up-to-date]	[How relevant a data asset is for the specific demo site. Ranked 1 (Low) - 5 (High)]	[How critical a data asset is for the demonstrator. Ranked 1 (Low) - 5 (High)]	[Explain the reasons for importance ranking]	
Energy production & delivery system:																																
EM1.1	Energiamittari	Kamstrup MC603 UF15 DN50	24-02-21			Float	csv	Real-Time			LJH = heat distribution room	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	4		
EM1.2	Energiamittari	Kamstrup MC603 UF3.5 DN25	24-02-21			Float	csv	Real-Time			LJH	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	4		
EM2.1	Energiamittari	Kamstrup MC603 UF15 DN50	24-02-21			Float	csv	Real-Time			LJH	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	4		
EM2.2	Energiamittari	Kamstrup MC603 UF3.5 DN25	24-02-21			Float	csv	Real-Time			LJH	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	4		
EM4.1	Energiamittari	Kamstrup MC603 UF25 DN65	24-02-21			Float	csv	Real-Time			LJH	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	4		
EM4.2	Energiamittari	Kamstrup MC603 UF25 DN65	24-02-21			Float	csv	Real-Time			LJH	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	4		
EM4.3	Energiamittari	Kamstrup MC603 UF30 DN40	24-02-21			Float	csv	Real-Time			LJH	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	4		
PDE			24-02-21			Float	csv	Real-Time				Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
PE4	PAINELÄHETIN	HK Instruments PTL-Cool-6-V	24-02-21			Float	csv	Real-Time			LJH	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
PE1	PAINELÄHETIN	HK Instruments PTL-Heat-6-V	24-02-21			Float	csv	Real-Time			LJH	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE1.1	LÄMPÖTILA-ANTURI	Produal TEAT NTC 10-80	24-02-21			Float	csv	Real-Time			LJH	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE1.2	LÄMPÖTILA-ANTURI	Produal TEAT NTC 10-50	24-02-21			Float	csv	Real-Time			LJH	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE1.3	LÄMPÖTILA-ANTURI	Produal TENA NTC 10-C	24-02-21			Float	csv	Real-Time			LJH	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE1.4	LÄMPÖTILA-ANTURI	Produal TENA NTC 10-C	24-02-21			Float	csv	Real-Time			LJH	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE1.5	LÄMPÖTILA-ANTURI	Produal TEAT NTC 10-80	24-02-21			Float	csv	Real-Time			LJH	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE1.6	LÄMPÖTILA-ANTURI	Produal TEAT NTC 10-80	24-02-21			Float	csv	Real-Time			LJH	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE2	LÄMPÖTILA-ANTURI	Produal TEU NTC 10	24-02-21			Float	csv	Real-Time			Ulkosäiä = external wall	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE3.1	LÄMPÖTILA-ANTURI		24-02-21			Float	csv	Real-Time			LVV1.1 = DHW storage	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE3.2	LÄMPÖTILA-ANTURI		24-02-21			Float	csv	Real-Time			LVV1.2	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE3.3	LÄMPÖTILA-ANTURI		24-02-21			Float	csv	Real-Time			LVV2.1	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE3.4	LÄMPÖTILA-ANTURI		24-02-21			Float	csv	Real-Time			LVV2.2	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE3.5	LÄMPÖTILA-ANTURI	Produal TENA NTC 10-C	24-02-21			Float	csv	Real-Time			LV = hot water	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE3.6	LÄMPÖTILA-ANTURI	Produal TENA NTC 10-C	24-02-21			Float	csv	Real-Time			LJH	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE3.7	LÄMPÖTILA-ANTURI	Produal TENA NTC 10-C	24-02-21			Float	csv	Real-Time			LV	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE3.8	LÄMPÖTILA-ANTURI	Produal TENA NTC 10-C	24-02-21			Float	csv	Real-Time			LVK	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE4.11	LÄMPÖTILA-ANTURI	TF25 NTC10k T150 050.06 5m	24-02-21			Float	csv	Real-Time			LJH	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE4.12	LÄMPÖTILA-ANTURI	TF25 NTC10k T150 050.06 5m	24-02-21			Float	csv	Real-Time			LJH	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE4.21	LÄMPÖTILA-ANTURI	TF25 NTC10k T150 050.06 5m	24-02-21			Float	csv	Real-Time			LJH	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE4.22	LÄMPÖTILA-ANTURI	TF25 NTC10k T150 050.06 5m	24-02-21			Float	csv	Real-Time			LJH	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE4.23	LÄMPÖTILA-ANTURI	TF25 NTC10k T150 050.06 5m	24-02-21			Float	csv	Real-Time			LJH	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE4.24	LÄMPÖTILA-ANTURI	TF25 NTC10k T150 050.06 5m	24-02-21			Float	csv	Real-Time			LJH	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE4.31	LÄMPÖTILA-ANTURI	TF25 NTC10k T150 050.06 5m	24-02-21			Float	csv	Real-Time			LJH	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE4.32	LÄMPÖTILA-ANTURI	TF25 NTC10k T150 050.06 5m	24-02-21			Float	csv	Real-Time			LJH	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE4.33	LÄMPÖTILA-ANTURI	TF25 NTC10k T150 050.06 5m	24-02-21			Float	csv	Real-Time				Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE4.34	LÄMPÖTILA-ANTURI	TF25 NTC10k T150 050.06 5m	24-02-21			Float	csv	Real-Time				Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE4.41	LÄMPÖTILA-ANTURI	TF25 NTC10k T150 050.06 5m	24-02-21			Float	csv	Real-Time			Kaivot = energy well	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE4.42	LÄMPÖTILA-ANTURI	TF25 NTC10k T150 050.06 5m	24-02-21			Float	csv	Real-Time			Kaivot	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE4.43	LÄMPÖTILA-ANTURI	TF25 NTC10k T150 050.06 5m	24-02-21			Float	csv	Real-Time			Kaivot	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE4.44	LÄMPÖTILA-ANTURI	TF25 NTC10k T150 050.06 5m	24-02-21			Float	csv	Real-Time			Kaivot	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE4.45	LÄMPÖTILA-ANTURI	TF25 NTC10k T150 050.06 5m	24-02-21			Float	csv	Real-Time			Kaivot	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TE4.46	LÄMPÖTILA-ANTURI	TF25 NTC10k T150 050.06 5m	24-02-21			Float	csv	Real-Time			Kaivot	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
EHP1	Electricity, heat pump 1		24-02-21			Float	csv	Real-Time			Technical room	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
EHP2	Electricity, heat pump 2		24-02-21			Float	csv	Real-Time			Technical room	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
EF_I	Electricity, circulation pumps (total or separate i)		24-02-21			Float	csv	Real-Time			Technical room	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
Air conditioning system (not fully described, in planning phase), from AHU yy:																																
QV_yy	based on dp)		24-02-21			Float	csv	Real-Time			AHU, air handling unit, centr	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
EAC_yy	conditioning/AHU		24-02-21			Float	csv	Real-Time			AHU, air handling unit, centr	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TSA_yy	Supply air temperature		24-02-21			Float	csv	Real-Time			AHU, air handling unit, centr	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TEXT_yy	Exit air temperature		24-02-21			Float	csv	Real-Time			AHU, air handling unit, centr	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
TWASTE_yy	Waste air temperature		24-02-21			Float	csv	Real-Time			AHU, air handling unit, centr	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
Measurements in apartments, xx indicating apartment number (not fully described, in planning phase):																																
TR_xx	Room air temperature		24-02-21			Float	csv	Real-Time			Apartment	Finnish		Raw				BAS & TAS							Proprietary/Private	Y					5	5		
RH1_xx	Room air humidity		24-02-21			Float	csv	Real-Time			Apartment	Finnish		Raw				BAS & TAS							Proprietary/Private	Y					5	5		
CO2_xx	Room air CO2 level		24-02-21			Float	csv	Real-Time			Apartment	Finnish		Raw				BAS & TAS							Proprietary/Private	Y					5	5		
TSET_xx	Room air temp setpoint		24-02-21			Float	csv	Real-Time			Apartment	Finnish		Raw				BAS & TAS							Proprietary/Private	Y					5	5		
QA_xx	Air flow rate (direct or control value %)		24-02-21			Float	csv	Real-Time			Apartment	Finnish		Raw				BAS & TAS							Proprietary/Private	Y					5	5		
OCC_xx	Occupancy (if possible, CO2 indicating occupancy too?)		24-02-21			Boolean	csv	Real-Time			Apartment	Finnish		Raw				BAS & TAS							Proprietary/Private	Y					5	5		
WFLOW_xx	Water flow rate of apartment xx		24-02-21			Float	csv	Real-Time			Apartment or technical room	Finnish		Raw				BAS & TAS							Proprietary/Private	Y					5	5		
EA_xx	xx		24-02-21			Float	csv	Real-Time			Apartment or technical room	Finnish		Raw				BAS & TAS							Proprietary/Private	Y					5	5		
General spaces of housing co-operative (not fully described):			24-02-21			Float	csv	Real-Time			Kaivot	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
EOUT1	section 1		24-02-21			Float	csv	Real-Time			Kaivot	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
EIN1	section 1		24-02-21			Float	csv	Real-Time			Kaivot	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
EOUT2	section 2		24-02-21			Float	csv	Real-Time			Kaivot	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
EIN2	section 2		24-02-21			Float	csv	Real-Time			Kaivot	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
EPARK	halls		24-02-21			Float	csv	Real-Time			Kaivot	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
ECAR	charging		24-02-21			Float	csv	Real-Time			Kaivot	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
Weather station data (not fully described):																																
T_OUT	Outdoor temperature		24-02-21			Float	csv	Real-Time			Weather station	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
RH3_OUT	Outdoor humidity		24-02-21			Float	csv	Real-Time			Weather station	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
SOLH	Solar radiation, horizontal		24-02-21			Float	csv	Real-Time			Weather station	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
SOLD	Solar radiation, diffuse		24-02-21			Float	csv	Real-Time			Weather station	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
SOLWEST	west	optional	24-02-21			Float	csv	Real-Time			Weather station	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
SOLSOUTH	south	optional	24-02-21			Float	csv	Real-Time			Weather station	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
WINDS	Wind speed		24-02-21			Float	csv	Real-Time			Weather station	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		
WINDD	Wind direction		24-02-21			Float	csv	Real-Time			Weather station	Finnish		Raw				BAS & TAS							Proprietary/Private	N					5	5		

Spanish demo site:

General Info				Data Asset Features													Data Asset Availability			Data Asset Rights		Data Asset Assessment		
ID	Data Source	Description	Datasets Available	Volume	Variety	Type	Format	Velocity	Historical Data Availability	Temporal Coverage	Spacial Coverage	Language	Relevant Standards	Veracity	Granularity	Dependency / Linking to Other Sources	Accessibility Method	Frequency of Updates	Documentation	Privacy	License	Accuracy	Completeness	Timeliness
[Unique identifier following the convention "OPEN_itno"]	The source (organisation) that provides the data	A brief description of the data asset	[Unique identifier following the convention "ODS_Organisation name_itno"]	[X GBs / records / transactions per hour / day / month]	[Structured / Unstructured / Semi-structured]	[Text / Image / Other]	[csv, xml, json, other]	[Real-time, Near Real-time, Batch]	[Y/N]	[From ... To...]	[Locations]	[e.g. Spanish, Finnish, French, Dutch, German, English...]	[List the international standards to which a data asset complies]	[Raw, Pre-processed, Processed Data asset]	[The temporal "granularity" of the data, e.g. per minute / hour / day / month]	[Y/N, If Y, list the other sources or codelists]	[Through API, As downloadable file, As database extract, Other]	[Real-time, Every X minutes / hours, Daily, Weekly, Monthly, Yearly, other]	The documentation of the API or data sample (incl. the location and the name of the file in the EXCESS repository)	[Confidential (not to be shared at all) / Proprietary (to be shared with appropriate licensing with the demonstrator partners) / Private (to be shared with appropriate licensing within the demonstrator / Public (available to all)]	[Exact Licence that is currently applied, e.g. CC Attribution-NonCommercial-ShareAlike (CC BY-NC-SA), or Case-by-Case Bilateral Agreement]	[Measure of correctness and precision, e.g. whether the dataset is effor-free, Ranked 1 (Low) - 5 (High)]	[Degree to which a data asset is sufficient in scope, depth, Ranked 1 (Low) - 5 (High)]	[How long a data asset remains up-to-date]
OPEN_1 (they are not open actually)	CEN	Exterior Temperature	ODS_CEN_1	1 / hour	Structured	Text	csv	Batch	Y			Spanish/English			Processed	hour		API	horly		Private		5	5	End-of project
OPEN_2	CEN	Solar Radiance on Horizontal	ODS_CEN_2	1 / hour	Structured	Text	csv	Batch	Y			Spanish/English			Processed	hour		API	horly		Private		5	5	End-of project
OPEN_3	Users	Aggregated electricity consumption in dwellings	ODS_CEN_3	1 / hour	Structured	Text	csv	Batch	Y			Spanish/English			Processed	hour		API	horly		Private		5	5	End-of project
OPEN_4	Users	Aggregated HVAC consumption in dwellings	ODS_CEN_4	1 / hour	Structured	Text	csv	Batch	Y			Spanish/English			Processed	hour		API	horly		Private		5	5	End-of project
OPEN_5	CEN	PV generation electricity	ODS_CEN_5	1 / hour	Structured	Text	csv	Batch	Y			Spanish/English			Processed	hour		API	horly		Private		5	5	End-of project
OPEN_6	CEN	Electricity bought from the grid	ODS_CEN_6	1 / hour	Structured	Text	csv	Batch	Y			Spanish/English			Processed	hour		API	horly		Private		5	5	End-of project
OPEN_7	CEN	Electricity sold to the grid	ODS_CEN_7	1 / hour	Structured	Text	csv	Batch	Y			Spanish/English			Processed	hour		API	horly		Private		5	5	End-of project
OPEN_8	CEN	Self-Consumption ratio	ODS_CEN_8	1 / hour	Structured	Text	csv	Batch	Y			Spanish/English			Processed	hour		API	horly		Private		5	5	End-of project
OPEN_9	CEN	Electricity consumption in Heat Pumps	ODS_CEN_9	1 / hour	Structured	Text	csv	Batch	Y			Spanish/English			Processed	hour		API	horly		Private		5	5	End-of project
OPEN_10	CEN	Thermal production consumption in Heat Pumps	ODS_CEN_10	1 / hour	Structured	Text	csv	Batch	Y			Spanish/English			Processed	hour		API	horly		Private		5	5	End-of project
OPEN_11	CEN	Electricity consumption in Pumps	ODS_CEN_11	1 / hour	Structured	Text	csv	Batch	Y			Spanish/English			Processed	hour		API	horly		Private		5	5	End-of project
OPEN_12	CEN	Electricity consumption in Electric Vehicles Charg	ODS_CEN_12	1 / hour	Structured	Text	csv	Batch	Y			Spanish/English			Processed	hour		API	horly		Private		5	5	End-of project
OPEN_13	CEN	Electric consumption in common zones	ODS_CEN_13	1 / hour	Structured	Text	csv	Batch	Y			Spanish/English			Processed	hour		API	horly		Private		5	5	End-of project
OPEN_14	CEN	CO2 emissions (conversion from kWh from grid a	ODS_CEN_14	1 / hour	Structured	Text	csv	Batch	Y			Spanish/English			Processed	hour		API	horly		Private		5	5	End-of project
OPEN_15	CEN	Aggregated Energy Balace (PEB definition)	ODS_CEN_15	1 / hour	Structured	Text	csv	Batch	Y			Spanish/English			Processed	hour		API	horly		Private		5	5	End-of project
OPEN_16	CEN	Flexibility	ODS_CEN_16	1 / hour	Structured	Text	csv	Batch	Y			Spanish/English			Processed	hour		API	horly		Private		5	5	End-of project

Austrian Demo site:

Data Description				Data Asset Features														Data Asset Availability							Data Asset Rights		Data Analysis	Data Asset Assessment					
Dataset ID	Data Asset Title	Description	Date of Row Last Update	Volume	Variety	Type	Format	Velocity	Historical Data Availability	Temporal Coverage	Spatial Coverage	Language	Relevant Standards	Veracity	Temporal Resolution	Spatial Resolution	Dependency / Linking to Other Sources	Data Asset Owner	Data Asset Available from 3rd Party	Data Asset Provider	Accessibility Method	Frequency of Updates	Update Strategy	Documentation	Privacy	Need for Anonymization	Types of analysis currently conducted on data	Accuracy	Completeness	Timeliness	Relevance to specific demonstration activities	Importance	Rationale	Comments
[Unique identifier following the convention "Country_Partner_DA_itno"]	The title of the data asset	A brief description of the data asset - At least 2-3 lines to give an overview of the data	The date of the last update of this row in the specific spreadsheet	[X GBs / records / transactions per hour / day / month]	[Structured / Unstructured / Semi-structured]	[Text / Image / Other]	[csv, xml, json, other]	[Real-time, Near Real-time, Batch]	[Y/N]	[From ... To...]	[Locations]	[e.g. Spanish, Finnish, French, Dutch, German, English...]	[List the international standards to which a data asset complies]	[Raw, Pre-processed, Processed Data asset]	[The temporal "granularity" of the data, e.g. per minute / hour / day / month]	[The spatial "granularity" of the data, e.g. at district / zone / building / area level]	[Y/N, If Y, list the other sources or codelists]	The name of the data asset owner	[Y/N]	The name of the data asset provider in EXCESS	[Through API, As downloadable file, As database extract, Other]	[Real-time, Every X minutes / hours, Daily, Weekly, Monthly, Yearly, other]	[Append/ newsfeed/ Replace-existing data / other]	The documentation of the API or data sample (incl. the location and the name of the file in the EXCESS repository)	[Confidential (not to be shared at all) / Proprietary (to be shared with appropriate licensing with the demonstrator partners) / Private (to be shared with appropriate licensing within the demonstrator / Public (available to all)]	[Y/N depending on whether the data asset contains sensitive or personal data]	[e.g. Correlation analysis for..., Predictive analytics for...]	[Measure of correctness and precision, e.g. whether the dataset is error-free, Ranked 1 (Low) - 5 (High)]	[Degree to which a data asset is sufficient in scope, depth, Ranked 1 (Low) - 5 (High)]	[How long a data asset remains up-to-date]	[How relevant a data asset is for the specific demo site, Ranked 1 (Low) - 5 (High)]	[How critical a data asset is for the demonstrator, Ranked 1 (Low) - 5 (High)]	[Explain the reasons for importance ranking]	
CYPRUS_SS_DA_1	Test dataset title	This is a description	20-01-2021	100 MB																													
P17-1010101-00-T01	TEMPLATE		19-02-2021	records	structured	Float	JSON	Real-Time	Y	from: set up time; to: current	ID: PPP-XXFFYZZ-BB-TAA PPP... Project/d, XX..Cluj/s, FF: floor-10, Y:dwelling Nr, ZZ..Room Nr (+10 subterrean), BB.. Category, T..Type of Sensor, AA.Nr of Sensor	English	RestfulAPI	Raw	12 per hour	room/machine	N	BAR	n	TSI	API	Real-time	Append		Private	N	raw							
P17-...-TAA	Site based data TEMPLATE	T ... type of sensor, AA ... identifier of sensor	19-02-2021	records	structured							English	RestfulAPI	Raw		site	N	BAR	y	openmeasherre	API	Real-time	Append										
P17-...-T02	Site Temp	site based (P17) local weather data current Temperature	19-02-2021	records	structured	Float	JSON	Real-Time			site	English	RestfulAPI	Raw		site	N	BAR	y	openmeasherre	API	Real-time	Append										
P17-...-T02	Site ForeCast Temp	site based (P17) local weather data forecast Temperature	19-02-2021	records	structured	Float	JSON	Real-Time			site	English	RestfulAPI	Raw		site	N	BAR	y	openmeasherre	API	Real-time	Append										
											0																							
P17-10-BB-TAA	Building based data TEMPLATE	BB ... category/machine identifier, T ... type of sensor, AA ... identifier of	19-02-2021	records	structured						building					building																	
P17-10-00-H01	Building Heating	building (P17-10) heating and cooling power consumption	19-02-2021	records	structured	Float	JSON	Real-Time	Y	from: set up time; to: current	building	English	RestfulAPI	Raw	12 per hour	building	N	BAR	n	TSI	API	Real-time	Append		Private	N	raw	4	5	5	5	5 occupant's comfort		
P17-10-00-E01	Building Electricity	building (P17-10) electrical power consumption	19-02-2021	records	structured	Float	JSON	Real-Time	Y	from: set up time; to: current	building	English	RestfulAPI	Raw	12 per hour	building	N	BAR	n	TSI	API	Real-time	Append		Private	N	raw	4	5	5	5	5 energy optimization		
P17-10-00-P01	Building Electricity generation	building (P17-10) electrical power production (photo voltaic)	19-02-2021	records	structured	Float	JSON	Real-Time	Y	from: set up time; to: current	building	English	RestfulAPI	Raw	12 per hour	building	N	BAR	n	TSI	API	Real-time	Append		Private	N	raw	4	5	5	5	5 energy optimization		
											0																							
P17-10FF-BB-TAA	dwelling based data TEMPLATE x10	FF ... floor (FF >= 10: ground level and above-ground floors; FF < 10: s	19-02-2021	records							dwelling					dwelling																	
P17-10101-00-H01	Dwelling Heating	dwelling (P17-10101) heating and cooling power consumption	19-02-2021	records	structured	Float	JSON	Real-Time	Y	from: set up time; to: current	dwelling	English	RestfulAPI	Raw	12 per hour	dwelling	N	BAR	n	TSI	API	Real-time	Append		Private	N	raw	4	5	5	5	5 occupant's comfort		
P17-10101-00-E01	Dwelling Electricity	dwelling (P17-10101) electrical power consumption	19-02-2021	records	structured	Float	JSON	Real-Time	Y	from: set up time; to: current	dwelling	English	RestfulAPI	Raw	12 per hour	dwelling	N	BAR	n	TSI	API	Real-time	Append		Private	N	raw	4	5	5	5	5 energy optimization		
P17-10101-00-X01	Dwelling Occupancy	dwelling (P17-10101) occupancy	19-02-2021	records	structured	bool	JSON	Real-Time	Y	from: set up time; to: current	dwelling	English	RestfulAPI	Raw	on change	dwelling	N	BAR	n	TSI	API	Real-time	Append		Private	Y	raw	2	2	2	2	2 energy optimizat	Opt-out, occupa	
											0																							
P17-10FFYZZ-BB-TAA	based data TEMPLATE 1 - 5 per DwFF	floor (FF >= 10: ground level and above-ground floors; FF < 10: s	19-02-2021	records	structured						room					room																	
P17-1010101-00-T01	Room Temperature	room temperature	19-02-2021	records	structured	Float	JSON	Real-Time	Y	from: set up time; to: current	room	English	RestfulAPI	Raw	12 per hour	room	N	BAR	n	TSI	API	Real-time	Append		Private	N	raw	4	5	5	5	5 occupant's comfort		
P17-1010101-00-F01	Room Humidity	room humidity	19-02-2021	records	structured	Float	JSON	Real-Time	Y	from: set up time; to: current	room	English	RestfulAPI	Raw	12 per hour	room	N	BAR	n	TSI	API	Real-time	Append		Private	N	raw	4	5	5	5	4 occupant's comfort		
P17-1010101-00-L01	Room Luminance	room luminance	19-02-2021	records	structured	Float	JSON	Real-Time	Y	from: set up time; to: current	room	English	RestfulAPI	Raw	12 per hour	room	N	BAR	n	TSI	API	Real-time	Append		Private	N	raw	4	5	5	5	4 occupant's comfort		
P17-1010101-00-C01	Room CO2 Concentration	room CO2 concentration	19-02-2021	records	structured	Float	JSON	Real-Time	Y	from: set up time; to: current	room	English	RestfulAPI	Raw	12 per hour	room	N	BAR	n	TSI	API	Real-time	Append		Private	N	raw	4	5	5	5	4 occupant's comfort		
P17-1010101-00-H01	Room Heating	room heating and cooling power (calculated)	19-02-2021	records	structured	Float	JSON	Real-Time	Y	from: set up time; to: current	room	English	RestfulAPI	Processed	12 per hour	room	linked with com	BAR	n	TSI	API	Real-time	Append		Private	N	calculated based on re	4	5	5	5	5 occupant's comfort		
P17-1010101-00-S01	Room Shading	room shading	19-02-2021	records	structured	Float	JSON	Real-Time	Y	from: set up time; to: current	room	English	RestfulAPI	Raw	on change	room	N	BAR	n	TSI	API	Real-time	Append		Private	N	raw	4	5	5	4	4 occupant's comfort		
P17-1010101-00-G01	Room Temperature Setpoint	room temperature setpoint	19-02-2021	records	structured	Float	JSON	Real-Time	Y	from: set up time; to: current	room	English	RestfulAPI	Raw	on change	room	N	BAR	n	TSI	API	Real-time	Append		Private	N	raw	5	5	5	5	5 occupant's comfort		

Belgian demo site:

Data Description		Data Description		Data Asset Features														Data Asset Availability								Data Asset Rights		Data Analysis	Data Asset Assessment					
Dataset ID	Data Asset Title	Description	Date of Row Last Update	Volume	Variety	Type	Format	Velocity	Historical Data Availability	Temporal Coverage	Spatial Coverage	Language	Relevant Standards	Veracity	Temporal Resolution	Spatial Resolution	Dependency / Linking to Other Sources	Data Asset Owner	Data Asset Available from 3rd Party	Data Asset Provider	Accessibility Method	Frequency of Updates	Update Strategy	Documentation	Privacy	Need for Anonymization	Types of analysis currently conducted on data	Accuracy	Completeness	Timeliness	Relevance to specific demonstration activities	Importance	Rationale	Comments
[Unique identifier following the convention <Country_Partner_DA_ incr]	The title of the data asset	A brief description of the data asset - At least 2-3 lines to give an overview of the data	The date of the last update of this row in the specific spreadsheet	[X GBs / records / transactions per hour / day / month]	[Structured / Semi-structured / Semi-structured]	[Text / Image / Other]	[csv, xml, json, other]	[Real-time, Near Real-time, Batch]	[Y/N]	[From … To…]	[Locations]	[e.g. Spanish, French, Dutch, German, English…]	[List the international standards to which a data asset complies]	[Raw, Pre-processed, Aggregated data asset]	[The temporal "granularity" of the data, e.g. per minute / hour / day / month]	[The spatial "granularity" of the data, e.g. at district / building / area level]	[Y/N, if Y, list the other sources or clusters]	The name of the data asset owner	[Y/N]	The name of the data asset provider (CKAN, EXCESS)	[Through API, As downloadable file, via electronic archive, Other]	[Real-time, Every X minutes / hours, Daily, Weekly, Monthly, Yearly, semi]	[Append new data / Append new data / Other]	[The documentation of the API or data sample (incl. the location and the name of the file in the EXCESS repository)]	[Confidential (not to be shared at all) / Proprietary (to be shared with appropriate licensing with the demonstrator partners) / Private (to be shared with appropriate licensing within the demonstrator) / Public (available to all)]	[Y/N depending on whether the data asset contains sensitive or personal data]	[e.g. Correlation analysis for…, Predictive analytics for…]	[Measure of correctness and precision, e.g. whether the dataset is error-free. Ranked 1 (Low) - 5 (High)]	[Degree to which a data asset is sufficient in scope, depth. Ranked 1 (Low) - 5 (High)]	[How/along a data asset remains up to-date]	[How relevant a data asset is for the specific demo-site. Ranked 1 (Low) -5 (High)]	[How critical a data asset is for the demonstrator. Ranked 1 (Low) -5 (High)]	[Explain the reasons for importance ranking]	
	District heating supply temperature	Supply temperature setpoint to the demo buildings	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	District heating return temperature	Return temperature setpoint to the demo buildings	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	District heating flowrate	District heating flowrate	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	District heating total heat demand	District heating total heat demand	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Central thermal energy storage tank SOC	State of charge of central storage tank	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Heat pump 1 condensor supply temperature	Heat pump 1 condensor supply temperature	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Heat pump 2 condensor supply temperature	Heat pump 2 condensor supply temperature	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Heat pump 3 condensor supply temperature	Heat pump 3 condensor supply temperature	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Heat pump 1 condensor return temperature	Heat pump 1 condensor return temperature	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Heat pump 2 condensor return temperature	Heat pump 2 condensor return temperature	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Heat pump 3 condensor return temperature	Heat pump 3 condensor return temperature	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Heat pump 1 evaporator supply temperature	Heat pump 1 evaporator supply temperature	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Heat pump 2 evaporator supply temperature	Heat pump 2 evaporator supply temperature	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Heat pump 3 evaporator supply temperature	Heat pump 3 evaporator supply temperature	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Heat pump 1 evaporator return temperature	Heat pump 1 evaporator return temperature	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Heat pump 2 evaporator return temperature	Heat pump 2 evaporator return temperature	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Heat pump 3 evaporator return temperature	Heat pump 3 evaporator return temperature	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Heat pump 1 condensor flowrate	Heat pump 1 condensor flowrate	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Heat pump 2 condensor flowrate	Heat pump 2 condensor flowrate	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Heat pump 3 condensor flowrate	Heat pump 3 condensor flowrate	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Heat pump 1 evaporator flowrate	Heat pump 1 evaporator flowrate	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Heat pump 2 evaporator flowrate	Heat pump 2 evaporator flowrate	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Heat pump 3 evaporator flowrate	Heat pump 3 evaporator flowrate	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Heat pump 1 thermal output	Heat pump 1 thermal output	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Heat pump 2 thermal output	Heat pump 2 thermal output	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Heat pump 3 thermal output	Heat pump 3 thermal output	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Heat pump 1 energy consumption	Heat pump 1 energy consumption	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Heat pump 2 energy consumption	Heat pump 2 energy consumption	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Heat pump 3 energy consumption	Heat pump 3 energy consumption	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	BTES supply temperature	Temperature of the fluid coming from the BTES	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	BTES return temperature	Temperature of the fluid entering the BTES	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	BTES flowrate	BTES flowrate	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	BTES thermal output	Heat output of the BTES	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Primary thermal buffer SOC	State of charge of cold storage tank	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	PVT supply temperature	Temperature of the fluid coming from the PVT panels	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	PVT return temperature	Temperature of the fluid entering the PVT panels	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	PVT flowrate	PVT flowrate	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	PVT thermal output	Heat output of the PVT	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	PVT electrical output	PVT electrical output	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Proprietary	N		5	5 min	5	5		
	Solar irradiation	Solar irradiation on the roof of the demo buildings	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Public	N		5	5 min	5	5		
	Outside temperature	Outside temperature at the demo buildings	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Public	N		5	5 min	5	5		
	Wind speed	Wind speed at the demo buildings	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Public	N		5	5 min	5	5		
	Wind direction	Wind direction at the demo buildings	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Area level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Public	N		5	5 min	5	5		
	Apartment 1…20 indoor temperature	Apartment 1…20 indoor temperature	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Dwelling level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Confidential / private	Y		5	5 min	5	5		
	Apartment 1…20 heat consumption	Apartment 1…20 heat consumption	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Dwelling level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Confidential / private	Y		5	5 min	5	5		
	Apartment 1…20 sensor x	Apartment 1…20 sensor x	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Dwelling level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Confidential / private	Y		5	5 min	5	5		
	…	…	17-02-21		Structured	Text, Float, …	json	Near Real-time	Y/N			English		Processed	5 min	Dwelling level	N	Cordium	Y	VITO	API, csv	5 min	Read only	TBD	Confidential / private	Y		5	5 min	5	5		